

30:1 BERKELEY TECHNOLOGY LAW JOURNAL

Spring 2015

Pages
1
to
384

Berkeley Technology Law Journal
Volume 30, Number 1

Production: Produced by members of the *Berkeley Technology Law Journal*.
All editing and layout done using Microsoft Word.

Printer: Joe Christensen, Inc., Lincoln, Nebraska.
Printed in the U.S.A.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Library Materials, ANSI Z39.48—1984.

Copyright © 2015 Regents of the University of California.
All Rights Reserved.



Berkeley Technology Law Journal
University of California
School of Law
3 Boalt Hall
Berkeley, California 94720-7200
btlj@law.berkeley.edu
<http://www.btlj.org>

BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 30

NUMBER 1

SPRING 2015

TABLE OF CONTENTS

ARTICLES

COMPELLING PASSWORDS FROM THIRD PARTIES: WHY THE FOURTH AND FIFTH AMENDMENTS DO NOT ADEQUATELY PROTECT INDIVIDUALS WHEN THIRD PARTIES ARE FORCED TO HAND OVER PASSWORDS.....	1
<i>Sarah Wilson</i>	
DISAGREEABLE PRIVACY POLICIES: MISMATCHES BETWEEN MEANING AND USERS' UNDERSTANDING	39
<i>Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, Roban Ramanath, N. Cameron Russell, Norman Sadeh & Florian Schaub</i>	
UNDERSTANDING THE FEDERAL CIRCUIT: AN EXPERT COMMUNITY APPROACH	89
<i>Laura G. Pedraza-Fariña</i>	
SOFTWARE COPYRIGHT'S ORACLE FROM THE CLOUD.....	161
<i>Lothar Determann & David Nimmer</i>	
SOVEREIGNTY UNDER SIEGE: CORPORATE CHALLENGES TO DOMESTIC INTELLECTUAL PROPERTY DECISIONS.....	213
<i>Cynthia M. Ho</i>	
IS COPYRIGHT INFRINGEMENT A STRICT LIABILITY TORT?	305
<i>Patrick R. Goold</i>	

SUBSCRIBER INFORMATION

The *Berkeley Technology Law Journal* (ISSN1086-3818), a continuation of the *High Technology Law Journal* effective Volume 11, is edited by the students of the University of California, Berkeley, School of Law (Boalt Hall) and is published in print three times each year (March, September, December), with a fourth issue published online only (July), by the Regents of the University of California, Berkeley. Periodicals Postage Rate Paid at Berkeley, CA 94704-9998, and at additional mailing offices. POSTMASTER: Send address changes to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94720-7210.

Correspondence. Address all correspondence regarding subscriptions, address changes, claims for non-receipt, single copies, advertising, and permission to reprint to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94705-7210; (510) 643-6600; JournalPublications@law.berkeley.edu. *Authors:* see section titled Information for Authors.

Subscriptions. Annual subscriptions are \$65.00 for individuals and \$85.00 for organizations. Single issues are \$30.00. Please allow two months for receipt of the first issue. Payment may be made by check, international money order, or credit card (MasterCard/Visa). Domestic claims for non-receipt of issues should be made within 90 days of the month of publication; overseas claims should be made within 180 days. Thereafter, the regular back issue rate (\$30.00) will be charged for replacement. Overseas delivery is not guaranteed.

Form. The text and citations in the *Journal* conform generally to the THE CHICAGO MANUAL OF STYLE (16th ed. 2010) and to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010). Please cite this issue of the *Berkeley Technology Law Journal* as 30 BERKELEY TECH. L.J. ____ (2015).

BTLJ ONLINE

The full text and abstracts of many previously published *Berkeley Technology Law Journal* articles can be found at <http://www.btlj.org>. Our site also contains a cumulative index, general information about the *Journal*, the *Bolt*, a collection of short comments and updates about new developments in law and technology written by BTLJ members, and *BTLJ Commentaries*, an exclusively online publication for pieces that are especially time-sensitive and shorter than typical law review articles.

INFORMATION FOR AUTHORS

The Editorial Board of the *Berkeley Technology Law Journal* invites the submission of unsolicited manuscripts. Submissions may include previously unpublished articles, essays, book reviews, case notes, or comments concerning any aspect of the relationship between technology and the law. If any portion of a manuscript has been previously published, the author should so indicate.

Format. Submissions are accepted in electronic format through the ExpressO online submission system. Authors should include a curriculum vitae and resume when submitting articles, including his or her full name, credentials, degrees earned, academic or professional affiliations, and citations to all previously published legal articles. The ExpressO submission website can be found at <http://law.bepress.com/expresso>.

Citations. All citations should conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010).

Copyrighted Material. If a manuscript contains any copyrighted table, chart, graph, illustration, photograph, or more than eight lines of text, the author must obtain written permission from the copyright holder for use of the material.

DONORS

The *Berkeley Technology Law Journal* and the Berkeley Center for Law & Technology acknowledge the following generous donors to Berkeley Law's Law and Technology Program:

Partners

COOLEY LLP

FENWICK & WEST LLP

COVINGTON & BURLING LLP

ORRICK, HERRINGTON &
SUTCLIFFE LLP

Benefactors

FISH & RICHARDSON P.C.

SKADDEN, ARPS, SLATE, MEAGHER
& FLOM LLP & AFFILIATES

KASOWITZ BENSON TORRES &
FRIEDMAN LLP

WEIL, GOTSHAL & MANGES LLP

KIRKLAND & ELLIS LLP

WHITE & CASE LLP

LATHAM & WATKINS LLP

WILMER CUTLER PICKERING HALE
AND DORR LLP

MCDERMOTT WILL & EMERY

WILSON SONSINI

MORRISON & FOERSTER LLP

GOODRICH & ROSATI

WINSTON & STRAWN LLP

Members

ALSTON & BIRD LLP	KEKER & VAN NEST LLP
BAKER BOTTS LLP	KILPATRICK TOWNSEND & STOCKTON LLP
BAKER & MCKENZIE LLP	KNOBBE MARTENS OLSON & BEAR LLP
BINGHAM MCCUTCHEN LLP	MUNGER, TOLLES & OLSON LLP
DURIE TANGRI LLP	O'MELVENY & MYERS LLP
GILLIN JACOBSON ELLIS LARSEN & LUCEY	PAUL HASTINGS LLP
GTC LAW GROUP LLP & AFFILIATES	ROPES & GRAY LLP
GUNDERSON DETTMER STOUGH VILLENEUVE FRANKLIN & HACHIGIAN, LLP	SIDLEY AUSTIN LLP
HAYNES AND BOONE, LLP	SIMPSON THACHER & BARTLETT LLP
HICKMAN PALERMO TRUONG BECKER BINGHAM WONG, LLP	TURNER BOYD LLP
HOGAN LOVELLS LLP	VAN PELT, YI & JAMES LLP
IRELL & MANELLA LLP	WEAVER AUSTIN VILLENEUVE & SAMPSON, LLP

BOARD OF EDITORS

2014–2015

Executive Committee

Editor-in-Chief

JOSEPH D. MORNIN

Managing Editor

EMILY LY

Senior Articles Editors

DALLAS BULLARD
MAISSA CHOURAKI
MARK ZAMBARDA

Senior Executive Editor

SIENA CARUSO

Senior Scholarship Editor

ANAGHA KRISHNAN

Senior Annual Review Editors

MISA EIRITZ
ALLISON SCHMITT

Senior Online Content Editor

ALI MAHANI

Editorial Board

Submissions Editors

CAROLINA GARCIA
ENEDA HOXHA

Production Editors

ROXANA GUIDERO
JENNA STOKES

Bluebook Editors

DIANA OBRADOVICH
SYDNEY RYAN
DANIELA SPENCER

Annual Review Editors

ALLISON CHO
LISA PATEL

Notes & Comments Editors

PRATUSHA ERRABALLI
THADDEUS HOUSTON

Symposium Editors

GINNY SCHOLTES
EMMIE TRAN

Commentaries Editors

NOAH DRAKE
KELLY VARGAS

Web & Technology Editors

LEIGHANNA MIXTER
STEVE SWANSON

External Relations Editor

JOHN RUSSELL

Member Relations Editor

NATHAN ANDERSEN

Bolt Editor

GOLDA CALONGE

Articles Editors

RAVI ANTANI
TOMMY BARCZYK
ELIZABETH CAPEL
SHELI DENOLA
ZACH FLOOD

SIMONE FRIEDLANDER
JOSH FURMAN
MEREDITH HOUSE
NICK HURT
VANESSA KATZ

JING LUO
JONATHAN MA
MATT REAGAN
MISHA TSUKERMAN
LILY WANG

MEMBERSHIP

Vol. 30 No. 1

Associate Editors

WAQAS AKMAL	AMARELLE HAYES	CHRISTELLE PRIDE
WILL BINKLEY	JEFF HIRSCHEY	LIDA RAMSEY
JOEL BROUSSARD	JENNIFER HSU	RACHAEL SHEN
MARIBETH CHARVET	HANNAH JIAM	MAXIMILIAN SLADEK DE LA CAL
STEPHANIE CHENG	CHARLES JOHNSON	BARTEK SUDOL
DAVID DERRICK	HILARY KRASE	JON TANAKA
DANEILLE DEVLIN	SARAH KWON	FAYE WHISTON
ERICA FISHER	JENNA LE	PHILIP WISEMAN
ZUBIN GAUTAM	NEETIKA MUTREJA	WINNIE WONG
YESOL HAN	ZACHARY NGUYEN	CHRISTOPHER YANDEL
CASSANDRA HAVENS	MICHAEL POON	
	SWAROOP POUDEL	

Members

YASMINE AGELIDIS	JINHYO LEE	ERIC RIEDEL
ISRAEL BER	STELLA LEE	ARIEL ROGERS
NORRIS BOOTHE	RICARDO LEYVA	SHINTARO SHIBA
MARIO BORISOV	BENJAMIN LI	BILLIE SOLOMON
JESSICA BRODSKY	MUQING LI	MARIA SOUZA
MARIA CANALES	XINYI LI	NATCHA SRIKAJON
GABRIELLE DORAIS	YU JU LIN	SAM SWARTZ
EVAN FERGUSON	PHILIP MERKSAMER	YU TANEBE
FELIPE FERNANDEZ	NICK MILLER-STRATTON	PAUL TISHER
ARI FINK	YAKOV MOLDAVSKIY	JANTHEN VAN
ANNA GASSOT	ELI NESS	RENTERGHEN
JAMES GATELY	PEGGY NI	KIRSTY WATKINS
MARK GRAY	THOMAS NOLAN	MEG WEBB
MAYA KAYE	CHRIS NORTON	JON WHITE
KAIA KIM	JORDYN OSTROFF	YUNPENG XIONG
MOTOYA KUBOTA	DYLAN PETERSON	SORIN ZAHARIA
CYNTHIA LEE	LEE REDFEARN	YU ZHAO
	JAMES RICE	MAYA ZIV

BTLJ ADVISORY BOARD

ROBERT BARR
*Executive Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

ROBERT C. BERRING, JR.
Walter Perry Johnson Professor of Law
U.C. Berkeley School of Law
Berkeley, California

JESSE H. CHOPER
Earl Warren Professor of Public Law
U.C. Berkeley School of Law
Berkeley, California

PETER S. MENELL
*Professor of Law and Faculty Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

ROBERT P. MERGES
*Wilson Sonsini Goodrich & Rosati Professor
of Law and Faculty Director of the Berkeley
Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

REGIS MCKENNA
Chairman and CEO
Regis McKenna, Inc.
Palo Alto, California

DEIRDRE K. MULLIGAN
*Assistant Professor and Faculty Director of the
Berkeley Center for Law and Technology*
U.C. Berkeley School of Information
Berkeley, California

JAMES POOLEY
*Deputy Director General of the
World Intellectual Property Organization*
Washington, D.C.

MATTHEW D. POWERS
Tensegrity Law Group, LLP
Redwood Shores, California

PAMELA SAMUELSON
*Professor of Law & Information
and Faculty Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

LIONEL S. SOBEL
Visiting Professor of Law
U.C.L.A. School of Law
Los Angeles, California

LARRY W. SONSINI
Wilson Sonsini Goodrich & Rosati
Palo Alto, California

MICHAEL STERN
Cooley LLP
Palo Alto, California

MICHAEL TRAYNOR
Cobalt LLP
Berkeley, California

THOMAS F. VILLENEUVE
Gunderson Dettmer Stough Villeneuve
Franklin & Hachigian LLP
Redwood City, California

BERKELEY CENTER FOR LAW & TECHNOLOGY 2014–2015

Executive Director

ROBERT BARR

Faculty Directors

KENNETH A. BAMBERGER
PETER S. MENELL

ROBERT P. MERGES
DEIRDRE MULLIGAN
PAMELA SAMUELSON

PAUL SCHWARTZ
MOLLY S. VAN HOUWELING

Associate Director

LOUISE LEE

Affiliated Faculty

CHRIS JAY HOOFNAGLE

TALHA SYED

JENNIFER M. URBAN

In Memoriam

SUZANNE SCOTCHMER

1950–2014

COMPELLING PASSWORDS FROM THIRD PARTIES: WHY THE FOURTH AND FIFTH AMENDMENTS DO NOT ADEQUATELY PROTECT INDIVIDUALS WHEN THIRD PARTIES ARE FORCED TO HAND OVER PASSWORDS

Sarah Wilson[†]

ABSTRACT

In 2012, the FBI served a search warrant on Google when a suspect—a user of Google’s phone services—refused to answer any questions about his cellphone, or provide the agents with the password to unlock it. The search warrant compelled Google to hand over the password information and other identifying information for the cellphone (account log-in, password reset, and manufacturer default code), which Google refused to do. Google’s refusal implicates a host of issues regarding our current understanding of privacy and self-incrimination protections and concerns legal scholars with what will happen to these doctrines if the government can simply bypass an individual and obtain passwords from a third party. This Article only begins to scratch the surface of this complex debate by analyzing the extent to which the Fourth and Fifth Amendments protect individuals when the government forces third parties to hand over their passwords, and will illustrate why these amendments do not adequately protect individuals in these situations. With constantly evolving technology and almost daily reports of the government accessing electronic communications and communication records, the time is ripe for Congress to legislate the issue of the government compelling private information, such as passwords, from third parties.

© 2015 Sarah Wilson

[†] J.D., Northwestern University School of Law, 2014; B.A., International Relations Global Business, University of Southern California, 2008. I would like to thank the members of the *Berkeley Technology Law Journal*, Jason Marsico, and Karin Lee for their invaluable comments and edits in developing this article. Most importantly, I would like to thank Matt Hinton for his love, support, and encouragement.

TABLE OF CONTENTS

I.	INTRODUCTION	2
II.	HOW RAPIDLY CHANGING COMMUNICATIONS TECHNOLOGY COULD THREATEN INDIVIDUALS' PRIVACY RIGHTS	6
III.	THE FOURTH AMENDMENT & THE REASONABLE EXPECTATION OF PRIVACY	9
A.	REASONABLE EXPECTATION OF PRIVACY IN THE CONTEXT OF EVOLVING TECHNOLOGY	10
B.	THE FOURTH AMENDMENT AND THIRD PARTIES	14
1.	<i>The Third-Party Doctrine & Voluntary Revelation</i>	15
2.	<i>Third Parties and the Content versus Non-Content Context</i>	16
3.	<i>The Employer Context</i>	18
4.	<i>What about Passwords?</i>	20
IV.	ATTEMPTS AT PROTECTING PASSWORDS UNDER THE FIFTH AMENDMENT	24
A.	ESTABLISHING THE REQUIREMENTS TO INVOKE THE FIFTH AMENDMENT	24
B.	COMPELLING PASSWORDS FROM INDIVIDUALS: DOESN'T THAT VIOLATE THE FIFTH AMENDMENT?	25
C.	ADDITIONAL CIRCUMVENTIONS OF THE FIFTH AMENDMENT	29
V.	INADEQUACY OF THE ELECTRONIC COMMUNICATIONS PRIVACY ACT	30
VI.	COMPELLING PASSWORDS FROM THIRD PARTIES	32
A.	SOLUTIONS FOR COMBATING COMPULSION OF PASSWORDS FROM SERVICE PROVIDERS	33
B.	COMPELLING PASSWORDS FROM EMPLOYERS	35
VII.	CONCLUSION	37

I. INTRODUCTION

“A legal battle is brewing between technology companies and the U.S. government over whether law-enforcement agents have the right to obtain

passwords¹ to crack into smartphones of suspects.”² Less than two years ago, the government seized a cellphone of a parolee suspected of becoming a “telephone pimp,” but was unable to access the contents of the locked phone because the parolee would not divulge the passcode and forensic technicians were unable to otherwise gain access.³ Although the government had obtained the cellphone from the suspect, the suspect refused to allow the government “access to it or answer any further questions.”⁴ Without the Google email username and password, however, the government could not access the contents of the seized cellphone.⁵ The government likely realized—based on precedent—that obtaining the password from the suspect would have been a fruitless effort under the Fifth Amendment, and one in which it did not care to rest its case.⁶ However, stumped by the

1. A password is “a secret word or expression used by authorized persons to prove their right to access” or “a word or other string of characters, sometimes kept secret or confidential, that must be supplied by a user in order to gain full or partial access to a multiuser computer system or its data resources.” *Password*, DICTIONARY.COM, <http://dictionary.reference.com/browse/password/> (last visited Feb. 8, 2015). Unless otherwise stated, “passwords” will be used to describe any passwords, encryption keys, or other strings of numbers or text that are used to secure data on an electronic device.

2. Julia Angwin, *FBI vs. Google: The Battle to Unlock Phones*, WALL ST. J., Sept. 7, 2012, at B4. See also Cyrus Farivar, *Feds Want Apple’s Help to Defeat Encrypted Phones, New Legal Case Shows*, ARS TECHNICA (Dec. 1, 2014, 9:00 AM), <http://arstechnica.com/tech-policy/2014/12/feds-want-apples-help-to-defeat-encrypted-phones-new-legal-case-shows/> (explaining that the DOJ has invoked the All Writs Act to facilitate the unlocking of a phone and noting the concern from at least one legal commentator that the “government’s application raises troubling questions about the extent to which it can force companies to break the products they sell”). In covering stories like these, recent news articles have fueled the idea that the U.S. government is compelling major internet companies to divulge users’ stored passwords and encryption algorithms. See, e.g., Declan McCullagh, *Feds Tell Web Firms to Turn Over User Account Passwords*, CNET (July 25, 2013, 11:26 AM), http://news.cnet.com/8301-13578_3-57595529-38/feds-tell-web-firms-to-turn-over-user-account-passwords/ (explaining that an internet industry source confirmed that he has seen the government ask for passwords and that they push back. Also, an employee at a large Silicon Valley company “confirmed that it received requests from the federal government for stored passwords.”).

3. Angwin, *supra* note 2, at B4.

4. Affidavit for Search Warrant at 5, *In re Search of Google Inc.*, No. 3:12-mj-00882-NLS (S.D. Cal. Mar. 9, 2012).

5. *Id.* at 7.

6. See, e.g., *United States v. Kirschner*, 823 F. Supp. 2d 665 (E.D. Mich. 2010); *In re Boucher (Boucher II)*, No. 2:06-mj-91, 2009 WL 424718, at *1, *4 (D. Vt. Feb. 19, 2009) (directing the defendant to provide an unencrypted version of his hard drive where the government’s expert was unable to search a computer because of password-protection and where the subpoena had asked for “any passwords used or associated with the [laptop]”). It is worth noting that in *In re Boucher (Boucher I)*, No. 2:06-mj-91, 2007 WL 4246473 (D. Vt. Nov. 29, 2007), *rev’d*, No. 2:06-mj-91, 2009 WL 424718 (D. Vt. Feb. 19, 2009), the magistrate judge found that the act of producing the password was testimonial and privileged, and as a

“pattern lock” on the phone, and the prospect of causing a lock-out by repeated incorrect attempts at the password, the government sought a search warrant ordering Google to “provide . . . any and all means of gaining access, including login and password information, password reset, and/or manufacturer default code.”⁷ Thus, the government chose to bypass the suspect’s potential Fifth Amendment claim by attempting to obtain the device’s password from the suspect’s cellphone manufacturer.⁸ The government sidestepped this potential roadblock and in so doing raised an important question: what is left of digital self-incrimination protection when the government can circumvent the Fifth Amendment by obtaining personal passwords from a third party?

Although Google refused to unlock the phone or even turn over the requested information,⁹ the implications of the government’s request, combined with constant innovation and evolving technology, warrant a discussion and a call for an update to the laws to prevent such workarounds from occurring without an individual’s consent. Moreover, the use of encryption, while just starting to take hold in the cellphone context, “is likely to reach into virtually every aspect of our lives.”¹⁰ Thus, Congress and the courts must carefully consider Americans’ constitutional rights in deciding whether to allow government compulsion of passwords from users and from third-party providers.

This Article will focus on the current understanding of password protections in the digital age, will consider to what extent the Fourth and Fifth Amendments protect individuals when the government forces third parties to hand over their passwords, and will illustrate why these amendments do not adequately protect individuals in these situations. This

result, on appeal, the request for a password was revised to a request for an unencrypted version of the hard drive. *See Boucher II*, 2009 WL 424718, at *1.

7. Affidavit for Search Warrant, *supra* note 4, at 1, 7 (“A pattern lock is a modern type of password installed on electronic devices, typically cellular telephones. To unlock the device, a user must move a finger or stylus over the keypad touch screen in a precise pattern so as to trigger the previously coded un-locking mechanism. Entering repeated incorrect patterns will cause a lock-out, requiring a Google email login and password to override.”); *see also* Angwin, *supra* note 2, at B4.

8. *See* Affidavit for Search Warrant, *supra* note 4.

9. *See* Search and Seizure Warrant at 2, *In re Search of Google Inc.*, No. 3:12-mj-00882-NLS (S.D. Cal. Mar. 26, 2012) (noting that “no property was obtained as Google Legal refused to provide the requested information”); *see also* Angwin, *supra* note 2, at B4 (noting that Google’s refusal was unusual and controversial and “indicates how murky the legal standards are for new technologies such as smartphones”).

10. *Privacy in the Digital Age: Encryption and Mandatory Access: Hearing Before the Subcomm. on the Constitution, Federalism, & Prop. Rights of the Comm. on the Judiciary*, 105th Cong. 3 (1998) (statement of Sen. Russell D. Feingold), available at <http://www.loc.gov/law/find/hearings/pdf/00139296461.pdf>.

Article will first address, in Part II, the rapidly changing technology that warrants a dialogue about individuals' waning privacy protections. Next, in Parts III and IV, this Article will discuss the Fourth Amendment privacy protections afforded to technology, the role of the Fifth Amendment in protecting passwords compelled by the government, and the shortcomings of these amendments in protecting passwords from third-party disclosures.

By exploring traditional analyses of the Fourth and Fifth Amendments and considering their application when the government compels passwords or other data “locking” mechanisms from third parties, this Article will illustrate the new paradigm that individuals face in their potential loss of constitutional protections. Unlike most articles that address whether the Fifth Amendment prevents the government from forcing an *individual* to provide a password or encryption key to permit access to his or her digital files,¹¹ this Article explores the issues that arise when the government compels *third parties* to provide a password or encryption key. To adequately address this issue, Part V will consider the impact of the Electronic Communications Privacy Act (“ECPA”) and specifically one subpart of it, the Stored Communications Act (“SCA”), when certain third parties are compelled to hand over user’s information. Finally, in Part VI, this Article will focus on the differences between two groups of third parties—service providers¹² and employers—that could be approached for passwords and will detail why, despite the differences, service providers and employers should be treated similarly when forced to compel data, at least at the present time.¹³ In addressing the two categories of third parties, Part VI will also detail potential solutions for combatting the compulsion of passwords from third parties and will explain why Congress should amend the ECPA to balance

11. See, e.g., Susan W. Brenner, *Encryption, Smart Phones, and the Fifth Amendment*, 33 WHITTIER L. REV. 525, 527, 535 (2012); Joshua A. Engel, *Rethinking the Application of the Fifth Amendment to Passwords and Encryption in the Age of Cloud Computing*, 33 WHITTIER L. REV. 543, 556 (2012); John E. D. Larkin, *Compelled Production of Encrypted Data*, 14 VAND. J. ENT. & TECH. L. 253, 264–72 (2012); Marjorie A. Shields, Annotation, *Fifth Amendment Privilege Against Self-Incrimination as Applied to Compelled Disclosure of Password or Production of Otherwise Encrypted Electronically Stored Data*, 84 A.L.R. 6TH 251 (2013).

12. As used in this Article, “service providers” refers to phone manufacturers, email and telecom providers, and any third-party service where users store password-protected information.

13. It is possible that a time will come when “keystroke logging” policies become commonplace in the workplace and that employees know that their employers can make permanent records of the passwords they enter into employer-provided devices. Should this change take place, employees will no longer have a reasonable expectation of privacy in the passwords they type into an employer-owned device and, in such a case, employers should be treated differently from third-party service providers.

the needs of law enforcement with the privacy interests of users in this rapidly changing digital world.

II. HOW RAPIDLY CHANGING COMMUNICATIONS TECHNOLOGY COULD THREATEN INDIVIDUALS' PRIVACY RIGHTS

The first generation cellphone was, plainly, a mobile phone, albeit a heavy one that was considered a rich man's toy.¹⁴ Similar to a landline, the cellphone was simply meant for talking. The cellphone of today is far more than that: it is a pocket-sized computer with email, texting, calling, GPS, and web-surfing capabilities.¹⁵ In fact, some may even say that "the term 'cell phone' is itself misleading shorthand; many of these devices are in fact minicomputers that also happen to have the capacity to be used as a telephone."¹⁶ As the cellphone has evolved, so too has the number of Americans who own one. As of January 2014, 90% of American adults have a cellphone.¹⁷ Some 63% of cellphone owners use their phones to access the internet, 52% exchange emails, and 50% download apps.¹⁸ Over 80% of cellphone owners use their phones to send or receive text messages.¹⁹ And 49% "get directions, recommendations, and other location-based information."²⁰ Thus, to a law enforcement agency, cellphone passwords "can be the key to unlocking a larger trove of information such as emails, texts, calls and address lists."²¹

14. Marguerite Reardon, *Cell Phone Industry Celebrates Its 25th Birthday*, CNET (Oct. 13, 2008, 1:52 PM), <http://www.cnet.com/news/cell-phone-industry-celebrates-its-25th-birthday/>.

15. See, e.g., *United States v. Zavala*, 541 F.3d 562, 577 (5th Cir. 2008) ("A cell phone is similar to a personal computer that is carried on one's person . . ."); Kristin Judge, *Protecting Your Mobile Phone: Password Lock and GPS Tracking Apps Can Help*, ANN ARBOR NEWS (Mar. 1, 2012, 10:00 AM), <http://www.annarbor.com/lifestyles/protecting-your-mobile-phone-password-locks-and-gps-tracking-applications-can-help/> ("In the 90s, a cellular phone was just a phone. The phone of 2012 is now a pocket-sized computer. The information you have access to from your phone is a wonderful thing."); Arthur Pinkasovitch, *Cell Phone Evolution*, INVESTOPEDIA (Mar. 19, 2010), <http://www.investopedia.com/financial-edge/0310/cell-phone-evolution.aspx> (explaining that apps on modern cellphones "help users watch movies, choose restaurants, do online banking, provide medical reference material, trade stocks, lose weight, navigate directions, read barcodes and performs millions of other fun and useful features").

16. *Riley v. California*, 134 S. Ct. 2473, 2489 (2014).

17. *Mobile Technology Fact Sheet*, PEW RESEARCH CTR., <http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/> (last visited Jan. 3, 2015).

18. *Id.*

19. *Id.*

20. *Id.*

21. Angwin, *supra* note 2, at B4.

As cellphones and computers have become more sophisticated, users have been able to conceal the data stored within them with more complex and sophisticated methods. While users can create a password to “lock” the device, they can also further protect devices by using encryption,²² “which requires a password or device . . . to decrypt the data into readable form.”²³ With password or encryption protection, files, directories, and applications are protected from unauthorized access.²⁴ Moreover, with encryption, “[o]nly the person holding the password is able to gain access to the original plaintext.”²⁵ While a number of smartphone apps have been available for some time now for “encrypting particular types of files, such as emails (i.e., NitroDesk TouchDown), voice calls (i.e., RedPhone), and text messages (i.e., Cypher),”²⁶ full-disk encryption for smartphones is a relatively new feature,²⁷ but one that is increasingly used.²⁸ Full-disk encryption is a way to convert

22. “Encryption is a process of translating a message, called the Plaintext, into an encoded message, called the Ciphertext. This is usually accomplished using a secret Encryption Key and a cryptographic Cipher. Two basic types of Encryption are commonly used: Symmetric Encryption, where a single secret key is used for both encryption and decryption [and] Asymmetric Encryption, where a pair of keys is used—one for Encryption and the other for Decryption.” *Definition of Encryption*, HITACHI ID SYS. INC., <http://hitachi-id.com/concepts/encryption.html> (last visited Aug. 24, 2013).

23. 117 AM. JUR. TRIALS 193 § 5 (2012).

24. “Encryption holds the promise of providing all of us with the ability to protect data and communications from unlawful and unauthorized access, disclosure, and alteration.” *Privacy in the Digital Age: Encryption and Mandatory Access: Hearing Before the Subcomm. on the Constitution, Federalism, & Prop. Rights of the Comm. on the Judiciary*, 105th Cong. 18 (1998) (prepared statement of Robert S. Litt, Principal Assoc. Deputy Att’y Gen.), available at <http://www.loc.gov/law/find/hearings/pdf/00139296461.pdf>; see also *Encrypting Your Smartphone Data*, BLACKBERRY SECURITY FEATURE OVERVIEW, http://docs.blackberry.com/en/smartphone_users/deliverables/39933/1812723.jsp (last visited Aug. 16, 2013) (“When you turn on encryption in the security options, your smartphone encrypts data stored on your smartphone (for example, browser information, messages, tasks, and calendar entries), including data that your smartphone receives when it is locked. If potentially malicious users attempt to access your data directly from the internal smartphone hardware, they can’t decrypt and read the data without knowing your smartphone password.”).

25. Michael S. Mahoney, Comment, *Compelling the Production of Passwords: Government’s Ability to Compel the Production of Passwords Necessary to the Discovery of Encrypted Evidence in Criminal Proceedings, Merely a Choice of Words*, 6 T.M. COOLEY J. PRAC. & CLINICAL L. 83, 89 (2003).

26. Ryan Radia, *Why You Should Always Encrypt Your Smartphone*, ARS TECHNICA (Jan. 16, 2011, 10:00 PM), <http://arstechnica.com/gadgets/2011/01/why-you-should-always-encrypt-your-smartphone/2/>.

27. *Id.* (noting that the iPhone 3GS, released in June 2009, marked Apple’s “first serious attempt” at full-disk encryption—which allowed the phone to be wiped remotely in seconds—while Motorola has stated that at least two of its Android smartphones “will soon offer full encryption”).

28. See, e.g., Craig Timberg, *Newest Androids Will Join iPhones in Offering Default Encryption, Blocking Police*, WASH. POST BLOG (Sept. 18, 2014), <http://www.washingtonpost.com/blogs/>

Although technology has changed, privacy rights should not. “At one point, [the Fourth and Fifth Amendment] together barred government from ‘any forcible and compulsory extortion of a man’s own testimony or of his private papers to be used as evidence’” against him.³⁵ Unfortunately the point in time where a compulsion was deemed to be an invasion of constitutional liberty and security, as embodied in the Fourth and Fifth Amendments,³⁶ was short-lived.³⁷ Americans should be protected against government intrusion into their passwords,³⁸ but, as this Article will illustrate, under current law, they are not.

III. THE FOURTH AMENDMENT & THE REASONABLE EXPECTATION OF PRIVACY

The Fourth Amendment guarantees “[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures”³⁹ Its fundamental purpose is to “safeguard the privacy and security of individuals against arbitrary invasions by government officials.”⁴⁰ In order to prove a violation of the Fourth Amendment, one must be able to demonstrate that a “search” occurred, i.e., that the government infringed upon an individual’s legitimate expectation of privacy.⁴¹ But, that expectation of privacy must be reasonable from both a subjective and objective point of view. The standard requires “first, that a person exhibited an actual or subjective expectation of privacy and second,

34. Julian Sanchez, *Old Technopanic in New iBottles*, CATO (Sept. 23, 2014, 5:17 PM), <http://www.cato.org/blog/old-technopanic-new-ibottles> (pointing out “excellent security reasons not to mandate backdoors” and asking us to think if encryption is “really any worse than a system of pay phones that allow criminals to communicate without leaving *any* record for police to sift through after the fact”).

35. Bryce Clayton Newell, *Rethinking Reasonable Expectations of Privacy in Online Social Networks*, 17 RICH. J.L. & TECH. 12, 24 (2011) (quoting *Boyd v. United States*, 116 U.S. 616, 630 (1886)).

36. *Boyd v. United States*, 116 U.S. 616, 630 (1886) (“The principles laid down in this opinion affect the very essence of constitutional liberty and security . . . they apply to all invasions on the part of the government and its employes of the sanctity of a man’s home and the privacies of life.”).

37. A little over thirty years after the Court found that the Fourth and Fifth Amendment protected individuals from government invasions, the Court began eroding this position in *Olmstead v. United States* by not finding police wiretapping to be an unreasonable search. 277 U.S. 438, 466 (1928).

38. Leslie Harris & Grover Norquist, *Protecting Our Privacy in the Digital Age*, THE HILL’S CONG. BLOG (June 19, 2013, 12:00 PM), <http://thehill.com/blogs/congress-blog/civil-rights/305735-protecting-our-privacy-in-the-digital-age/>.

39. U.S. CONST. amend. IV.

40. *Camara v. San Francisco Mun. Ct.*, 387 U.S. 523, 528 (1967).

41. *See, e.g., Rakas v. Illinois*, 439 U.S. 128 (1978); 68 AM. JUR. 2D *Searches & Seizures* § 6 (2014).

everything on a hard drive, including the operating system, into an unreadable form until the proper key (i.e., password) is entered.²⁹ And while privacy advocates are rightfully rejoicing over recent moves by Apple and Google to expand the use of encryption on smartphones, law enforcement fears a certain downside to encryption, specifically that it will become impossible to collect evidence from smartphones, even with a search warrant.³⁰ The FBI has reported that “encryption has been used to conceal criminal activity and thwart law enforcement efforts to collect critical evidence needed to solve serious and often violent criminal activities.”³¹ And while criminal activity may be fueling the government’s request for third parties to fork over passcodes, information security concerns are causing phone manufacturers to create engineering solutions to remove backdoor access (an ability for manufacturers to unlock a device) to a phone’s passcode, such that it will no longer be feasible for the manufacturers to respond to government warrants—unless the data is stored in the server’s cloud.³² This move has been questioned by some³³ and touted by others.³⁴

the-switch/wp/2014/09/18/newest-androids-will-join-iphones-in-offering-default-encryption-blocking-police/ (“The next generation of Google Android’s operating system . . . will encrypt data by default.”). One legal scholar in particular “believe[s] we will see an increased use of encryption and other data-protection measures that will make it increasingly difficult, if not impossible, for officers to access the contents of a smart phone or other digital device by bypassing minimal, if any, security measures.” Brenner, *supra* note 11, at 533.

29. KAREN SCARFONE ET AL., GUIDE TO STORAGE ENCRYPTION TECHNOLOGIES FOR END USER DEVICES 3-1 to -2 (2007), available at <http://csrc.nist.gov/publications/nistpubs/800-111/SP800-111.pdf>.

30. Timberg, *supra* note 28.

31. *Privacy in the Digital Age: Encryption and Mandatory Access: Hearing Before the Subcomm. on the Constitution, Federalism, & Prop. Rights of the Comm. on the Judiciary*, 105th Cong. 4 (1998) (statement of Sen. Russell D. Feingold), available at <http://www.loc.gov/law/find/hearings/pdf/00139296461.pdf>.

32. Craig Timberg, *Apple Will No Longer Unlock Most iPhones, iPads for Police, Even with Search Warrants*, WASH. POST (Sept. 18, 2014), http://www.washingtonpost.com/business/technology/apple-will-no-longer-unlock-most-iphones-ipads-for-police-even-with-search-warrants/2014/09/17/2612af58-3ed2-11e4-b03f-de718edeb92f_story.html.

33. See, e.g., Orin Kerr, *Apple’s Dangerous Game*, WASH. POST (Sept. 19, 2014), <http://www.washingtonpost.com/news/volokh-conspiracy/wp/2014/09/19/apples-dangerous-game/> (finding “Apple’s new design very troubling”). Although, Orin Kerr subsequently changed his view from “very troubled” to “need more information.” Orin Kerr, *Apple’s Dangerous Game, Part 2: The Strongest Counterargument*, WASH. POST (Sept. 22, 2014), <http://www.washingtonpost.com/news/volokh-conspiracy/wp/2014/09/22/apples-dangerous-game-part-2-the-strongest-counterargument/>; see also Craig Timberg & Greg Miller, *FBI Blasts Apple, Google for Locking Police Out of Phones*, WASH. POST (Sept. 25, 2014), http://www.washingtonpost.com/business/technology/2014/09/25/68c4e08e-4344-11e4-9a15-137aa0153527_story.html (“FBI Director James B. Comey sharply criticized Apple and Google . . . for developing forms of smartphone encryption so secure that law enforcement officials cannot easily gain access to information stored on the devices.”).

that the expectation be one that society is prepared to recognize as reasonable.”⁴² If this standard is met, the government may not compel this information from an individual or a third party without first obtaining a warrant based on probable cause.⁴³

A. REASONABLE EXPECTATION OF PRIVACY IN THE CONTEXT OF EVOLVING TECHNOLOGY

Courts and government officials alike have struggled to apply the Fourth Amendment expectation of privacy test to electronic devices and technology (such as computers, cellphones, emails, the internet, and encryption), and have yet to create a clear standard to be applied in the context of evolving technology.⁴⁴ But, “the Fourth Amendment must keep pace with the inexorable march of technological process, or its guarantees will wither and perish.”⁴⁵

(1) Computers. The Fourth Amendment protection afforded to computers is not well defined, and is often determined by the circumstances surrounding the computer files and the extent to which the files are protected. Courts, however, tend to start with the underlying assumption that individuals have a reasonable expectation of privacy in files stored on their hard drives, as closed computer files resemble closed containers.⁴⁶ The expectation of privacy tends to disappear, though, when: (1) the search of a computer is conducted at the U.S. border;⁴⁷ (2) the search is on an employer-owned computer where the employer gave notice of its policy to monitor computer use;⁴⁸ (3) when an individual’s computer files are accessible to

42. 117 AM. JUR. TRIALS 193 § 2 (2012); *see also* Katz v. United States, 389 U.S. 347, 361 (1967) (Harlan, J., concurring); United States v. Ward, 561 F.3d 414, 414 (5th Cir. 2009).

43. United States v. Warshak, 631 F.3d 266, 288 (6th Cir. 2010).

44. *See* 3A CHARLES ALAN WRIGHT ET AL., PRACTICE & PROCEDURE § 663 (4th ed. 2013) (“One of the most difficult questions raised by the reasonable-expectation-of-privacy test for defining searches is how it applies in a world of digital communications.”).

45. United States v. Warshak, 631 F.3d at 285.

46. *See* United States v. Barth, 26 F. Supp. 2d 929, 936 (W.D. Tex. 1998) (holding that the “Fourth Amendment protection of closed computer files and hard drives is similar to the protection it affords a person’s closed containers and closed personal effects”); *see also* United States v. Crist, 627 F. Supp. 2d 575, 586 (M.D. Pa. 2008); United States v. Knoll, 16 F.3d 1313, 1320 (2d Cir. 1994) (holding that “[i]f the files were closed and their contents not apparent from the exterior, the reasonable expectation of privacy continued so long as the files had not been searched before contact with the government occurred”).

47. *See* United States v. Arnold, 533 F.3d 1003, 1007 (9th Cir. 2008) (confirming that “searches of closed containers and their contents can be conducted at the border without particularized suspicion under the Fourth Amendment”).

48. *See* United States v. Simons, 206 F.3d 392, 398 (4th Cir. 2000).

others via a peer-to-peer file sharing software;⁴⁹ or (4) because the computer is shared with a third party⁵⁰ (unless separate passwords and accounts are set-up between the parties⁵¹).

(2) Cellphones. There is little dispute that an individual has a subjective expectation of privacy in their cellphone because of the wealth of information stored on it.⁵² After all, “the uniquely sensitive nature of data on electronic devices carries with it a significant expectation of privacy.”⁵³ However, courts have long questioned whether there is an objective expectation of privacy in a cellphone. Although courts have been split on this question, the Supreme Court recently stated in dicta that “a warrant is generally required before [searching a cellphone].”⁵⁴ However the case itself and the Court’s holding was focused on requiring a warrant “when a cell phone is seized incident to arrest.”⁵⁵ Thus, the law is not yet settled and it remains to be seen to what extent other courts will follow the general principle espoused by the Supreme Court. And, following this monumental case in favor of increased privacy protections of cellphones, it is unclear what will come of the rulings where courts have allowed government officials to search an individual’s cellphone without a warrant, such as when the

49. *See* United States v. Stults, 575 F.3d 834, 843 (8th Cir. 2009) (holding that an individual had no reasonable expectation of privacy in files retrieved from his personal computer because he made his files accessible to others through file sharing); *see also* United States v. Conner, 521 Fed. App’x 493, 494 (6th Cir. 2013) (explaining that the “ability to download a file directly from another user’s personal computer is known as ‘peer-to-peer’ file sharing”). For more information about peer-to-peer file sharing, see *What You Need to Know About Peer-to-Peer File Sharing*, ZONEALARM (June 4, 2014), <http://www.zonealarm.com/blog/2014/06/what-you-need-to-know-about-peer-to-peer-file-sharing/> (explaining that peer-to-peer file sharing “is the process of sharing and transferring digital files from one computer to another”).

50. *See* United States v. Matlock, 415 U.S. 164, 171 n.7 (1974) (holding that mutual use of property gives the co-inhabitants the right to permit inspection of shared common property).

51. *See* Trulock v. Freeh, 275 F.3d 391, 403 (4th Cir. 2001) (finding that although a shared user of the computer in question “had authority to consent to a general search of the computer,” and the joint hard drive, her authority did not extend to the other user’s password-protected files).

52. *See, e.g.*, United States v. Zavala, 541 F.3d 562, 577 (5th Cir. 2008) (recognizing that “cell phones contain a wealth of private information,” and, as a result, individuals have a “reasonable expectation of privacy” in them); State v. Smith, 920 N.E.2d 949, 954–55 (Ohio 2009) (recognizing that “a person has a high expectation of privacy in a cell phone’s contents” and therefore a subjective expectation of privacy in cellphones because “modern cell phones are capable of storing a wealth of digitized information”).

53. United States v. Cotterman, 709 F.3d 952, 966 (9th Cir. 2013).

54. Riley v. California, 143 S. Ct. 2473, 2493 (2014).

55. *Id.* at 2493–94 (2014) (leaving open the possibility that “other case-specific exceptions may still justify a warrantless search of a particular phone”).

government was seeking non-content data or attempting to preserve destructible data.⁵⁶ However, it will be hard to ignore the Supreme Court's finding that "[m]odern cell phones . . . implicate privacy concerns far beyond those implicated by the search of a cigarette pack, a wallet, or a purse."⁵⁷ In fact, it may be that more courts will follow in step with others that have ruled that the government cannot search an individual's cellphone without a warrant,⁵⁸ such that individuals will soon hold an explicit objective expectation of privacy in cellphones. But, perhaps with such ambiguity, it will be left to the states to take action legislatively to expressly state that a search warrant is needed to seek information on an electronic device.⁵⁹

(3) Emails. Although government entities can utilize the Stored Communications Act—an Act that addresses voluntary and compelled disclosure of information by third parties—“to compel a service provider to disclose the contents of [electronic] communications,”⁶⁰ most courts recognize, nevertheless, that there is both a subjective and objective expectation of privacy in the contents of emails and thus afforded emails Fourth Amendment protection.⁶¹ Emails often contain the contents of an

56. *See, e.g., In re U.S. for Historical Cell Site Data*, 724 F.3d 600, 611 (5th Cir. 2013) (relying on the third-party doctrine to compel historical location data from cellphone service providers without a warrant, and implying that this data was not subject to a reasonable expectation of privacy); *United States v. Murphy*, 552 F.3d 405, 411 (4th Cir. 2009).

57. *Riley*, 143 S. Ct. at 2489.

58. *See, e.g., United States v. Wurie*, 728 F.3d 1, 14 (1st Cir. 2013) (internal citation omitted), *aff'd sub nom. Riley v. California*, 134 S. Ct. 2473 (2014); *United States v. Zavala*, 541 F.3d 562, 576–77 (5th Cir. 2008) (holding that individuals have a reasonable expectation of privacy regarding the information in a cellphone and the search of a suspect's cellphone without a warrant violated the Fourth Amendment); *Smallwood v. State*, 113 So. 3d 724, 735 (Fla. 2013) (holding that once the electronic, computer-like phone was removed from the arrestee, the officer was “constitutionally required to obtain a warrant before searching the contents of, and the data in” the cellphone); *State v. Smith*, 920 N.E.2d 949, 955 (Ohio 2009) (finding that “because a person has a high expectation of privacy in a cell phone's contents, police must then obtain a warrant before intruding into the phone's contents”). But, even in a case decided after *Riley*, a court has held that Fourth Amendment protections do not apply to cell site location information. *See United States v. Banks*, 2014 U.S. Dist. LEXIS 128906 (D. Kan. Sept. 15, 2014).

59. *See, e.g., MONT. CODE ANN. § 46-5-110* (2013) (noting that “a government entity may not obtain the location information of an electronic device without a search warrant” unless authorized by the user, there is a life-threatening situation, or the user reports a call for emergency services or that the device has been stolen).

60. *Warshak v. United States*, 532 F.3d 521, 523 (6th Cir. 2008).

61. *See, e.g., United States v. Warshak*, 631 F.3d 266, 284 (6th Cir. 2010); *In re Applications for Search Warrants for Info. Associated with Target Email Address*, No. 12-MJ-8119-DJW, 2012 WL 4383917, at *5 (D. Kan. Sept. 21, 2012) (holding that “an individual has a reasonable expectation of privacy in emails” and that the Fourth Amendment protections apply to email); *In re Matter of the Search of Info. Associated with [redacted]@mac.com that is Stored at Premises Controlled by Apple, Inc.*, No. 14-228

individual's business and personal life, and it is highly unlikely that given their often sensitive and sometimes damning contents, individuals "expect[] them to be made public, for people seldom unfurl their dirty laundry in plain view."⁶² Like the telephone user, an email user is "entitled to assume that the words he utters into [a device] will not be broadcast to the world."⁶³ And given the fundamental similarities between email and other traditional forms of communication, like the telephone and the letter, it would "defy common sense to afford emails lesser Fourth Amendment protection."⁶⁴ Thus, society is prepared to recognize an expectation of privacy in emails.

(4) The internet and encryption. Recently, the IRS "argued that anyone who used the internet had no reasonable expectation of privacy against governmental intrusion."⁶⁵ However, some legal scholars have argued that an individual should rightfully claim a reasonable expectation of privacy in the contents of internet pages and cloud computing when the individual safeguards those contents through reasonable concealment efforts, such as password protection and encryption.⁶⁶

(JMF), 2014 WL 945563, at *8 (D.D.C. Mar. 7, 2014) (noting the "obvious expectation of privacy e-mail account holders have in their communications"); *United States v. DiTomasso*, No. 14-CR-160 SAS, 2014 WL 5462467, at *9 (S.D.N.Y. Oct. 28, 2014) (holding that defendant had a reasonable expectation of privacy in e-mails but waived this Fourth Amendment protection by agreeing to AOL's terms of service).

62. *Warshak*, 631 F.3d at 284.

63. *Katz v. United States*, 389 U.S. 347, 352 (1967).

64. *Warshak*, 631 F.3d at 285–86.

65. Ray Bishop, *Protecting Our Privacy in the Digital Age*, CRITICAL DECISIONS (June 21, 2013), <http://critical-decisions.com/Index/2013/06/21/protecting-our-privacy-in-the-digital-age/>. The 2009 search warrant handbook from the IRS Criminal Tax Division's Office of Chief Counsel asserts that "the Fourth Amendment does not protect communications held in electronic storage, such as email messages stored on a server, because internet users do not have a reasonable expectation of privacy in such communications." IRS OFFICE OF CHIEF COUNSEL CRIMINAL TAX DIVISION, SEARCH WARRANT HANDBOOK 59 (2009), available at <https://www.aclu.org/national-security/search-warrant-handbook/>. See also Nathan Freed Wessler, *New Documents Suggest IRS Reads Emails Without a Warrant*, AM. CIVIL LIBERTIES UNION (Apr. 10, 2013), <https://www.aclu.org/blog/technology-and-liberty-national-security/new-documents-suggest-irs-reads-emails-without-warrant/> (claiming that "IRS Criminal Tax Division has long taken the position that the IRS can read your emails without a warrant—a practice that one appeals court has said violates the Fourth Amendment"). But see Harris & Norquist, *supra* note 38 (noting that when this policy was brought to light, the IRS backed off and said that "it would obtain a search warrant in all cases when seeking from an Internet service provider the content of email communications stored on behalf of customers").

66. See, e.g., WAYNE R. LAFAVE, 1 SEARCH AND SEIZURE: A TREATISE ON THE FOURTH AMENDMENT § 2.6(f) (4th ed. 2004); David A. Couillard, Note, *Defogging the Cloud: Applying Fourth Amendment Principles to Evolving Privacy Expectations in Cloud Computing*, 93 MINN. L. REV. 2205, 2232 (2009).

In 1998, before encryption and passwords became the status quo on all electronic devices, Senator Russ Feingold remarked that “the fundamental right to privacy . . . will be at least somewhat sacrificed”⁶⁷ if an encryption user is denied the right to keep his personal information completely private by the inability to purchase non-recoverable encryption (i.e., encryption that can only be broken by the user’s password, not by a backdoor). But, it is unclear how a recoverable encryption key or password (a backup or “backdoor” decryption capability) affects a user’s reasonable expectation of privacy. If a third party is compelled to hold another key or backdoor to a cellphone’s password, the user’s password may not be protected by the Fourth Amendment, as one can hardly find a reasonable expectation of privacy in recoverable encryption.⁶⁸ However, with non-recoverable encryption, users are able to keep their information completely confidential, and a court would likely find both subjective and objective expectations of privacy. It remains to be seen whether the Supreme Court agrees.

B. THE FOURTH AMENDMENT AND THIRD PARTIES

As an initial matter, the mere *ability* of a third-party intermediary to access the contents of a communication cannot be sufficient to extinguish a reasonable expectation of privacy.⁶⁹ Nor does the *right* of access.⁷⁰

In *Katz v. United States*, telephone companies had both the ability and right to monitor calls, yet the Supreme Court found a reasonable expectation of privacy during a telephone call.⁷¹ Likewise, letters and other sealed packages carry with them an expectation of privacy for the sender,⁷² despite the fact that they are handed over to numerous mail carriers and are subjected to the risk that any mail handler “could tear open the thin paper

67. *Privacy in the Digital Age: Encryption and Mandatory Access: Hearing Before the Subcomm. on the Constitution, Federalism, & Prop. Rights of the Comm. on the Judiciary*, 105th Cong. 4 (1998) (statement of Sen. Russell D. Feingold), available at <http://www.loc.gov/law/find/hearings/pdf/00139296461.pdf>.

68. *Id.*; but see *United States v. Taketa*, 923 F.2d 665, 673 (9th Cir. 1991) (noting that if a court allowed the existence of a master key to overcome the expectation of privacy, it would defeat the legitimate privacy interest of any hotel, office, or apartment occupant, which our courts are not prepared to do). Thus, the existence of a master key may not impact a court’s analysis of a “reasonable expectation of privacy.”

69. *Warshak*, 631 F.3d at 286; see also *Katz v. United States*, 389 U.S. 347, 353 (1967) (concluding that the government’s access to a telephone conversation did not destroy the reasonable expectation of privacy).

70. *Warshak*, 631 F.3d at 287.

71. *Katz*, 389 U.S. at 353, 359; see also *Smith v. Maryland*, 442 U.S. 735, 746–47 (1979) (Stewart, J., dissenting) (stating that the telephone conversation in *Katz* “must be electronically transmitted by telephone company equipment, and may be recorded or overheard by the use of other company equipment”).

72. *United States v. Jacobsen*, 446 U.S. 109, 114 (1984).

envelopes that separate the private words from the world outside.”⁷³ Even the degree of access that Internet Service Providers (“ISPs”) have over email accounts does not necessarily diminish a reasonable expectation of privacy in an individual’s email.⁷⁴

Conversely, there are times when individuals assume the risk that the third party to whom they convey information may subsequently reveal that information to the government. This most often occurs when the information is “voluntarily” turned over or “knowingly exposed” to a third party.⁷⁵

1. *The Third-Party Doctrine & Voluntary Revelation*

Under Supreme Court precedent, information that is voluntarily revealed to third parties does not warrant Fourth Amendment protection.⁷⁶ Instead, this voluntary act of disclosure invokes the third-party doctrine, which gives the government additional leeway in obtaining “so-called private” information. Essentially, the third-party doctrine “postulates that if the information in question has been voluntarily turned over to a third party, the individual seeking privacy protection no longer has a reasonable expectation of privacy.”⁷⁷

Smith v. Maryland is a notable case for this doctrine. Here the Court found that individuals lack a reasonable expectation of privacy in the numbers they dial on a telephone.⁷⁸ To use a phone, an individual must type in numbers to connect with another party, and such an act “voluntarily convey[s] numerical

73. *Warshak*, 631 F.3d at 285.

74. *Id.* at 287 (noting that the ISP subscriber agreement specifically stated that the ISP “may access and use individual Subscriber information,” but that the degree of access did not diminish the reasonable expectation of privacy in emails).

75. *See, e.g., Katz*, 389 U.S. at 351 (finding that “[w]hat a person knowingly exposes to the public, even in his own home or office, is not a subject of Fourth Amendment protection”); *Reporters Comm. for Freedom of Press v. Am. Tel. & Tel. Co.*, 593 F.2d 1030, 1043 (D.C. Cir. 1978) (“To the extent an individual knowingly exposes his activities to third parties, he surrenders Fourth Amendment protection”).

76. Orin S. Kerr, *A User’s Guide to the Stored Communications Act, and a Legislator’s Guide to Amending It*, 72 GEO. WASH. L. REV. 1208, 1210 (2004) (“The Supreme Court has repeatedly held, however, that the Fourth Amendment does not protect information revealed to third parties.”); *see also Smith*, 442 U.S. at 743–44 (1979); *United States v. Miller*, 425 U.S. 435, 443 (1976); *United States v. White*, 401 U.S. 745, 749, 754 (1971).

77. Lindsay M. Gladysz, Note, *Status Update: When Social Media Enters the Courtroom*, 7 I/S: J. L. & POL’Y FOR INFO. SOC’Y 688, 708 (2012); *see also* Orin S. Kerr, *The Case for the Third-Party Doctrine*, 107 MICH. L. REV. 561, 563 (2009).

78. *Smith*, 442 U.S. at 744–46. *But cf. In re Application of U.S. for an Order Directing a Provider of Elec. Comm’n Serv. to Disclose Records to the Gov’t*, 620 F.3d 304, 317 (3d Cir. 2010) (noting that although cellphone customers may voluntarily share phone numbers dialed, they do not voluntarily share cell site location information).

information to the telephone company”⁷⁹ as individuals must know that phone companies have “facilities for making permanent records of the numbers they dial”⁸⁰ As a result, a “person has no legitimate expectation of privacy in information he voluntarily turns over to third parties.”⁸¹ However, it is unlikely that cellphone customers are aware that their cellphone providers collect any additional information aside from numbers dialed into the phone.⁸² Thus, whether the expectation of privacy is “one that society is prepared to recognize as reasonable” is dependent, at least in part, on whether information is *voluntarily* conveyed and *knowingly* used.⁸³

2. *Third Parties and the Content versus Non-Content Context*

In analyzing cases involving third parties, another way to characterize the split is through the content versus non-content divide.⁸⁴ The distinction between *Smith* and *Katz* most prominently lies in the difference between non-content information (often thought of as metadata)⁸⁵ and content information.

In *Smith* and its subsequent line of cases, courts have upheld the notion that users do not have a “reasonable expectation of privacy in their subscriber information, the length of their stored files, and other *non-content* data to which service providers must have access”⁸⁶ (such as phone numbers

79. *Smith*, 442 U.S. at 744.

80. *Id.* at 742.

81. *Id.* at 743–44; *see also Miller*, 425 U.S. at 442 (holding that a bank depositor does not have a legitimate expectation of privacy in the “financial statements and deposits slips . . . voluntarily conveyed to the banks and exposed to their employees in the ordinary course of business”).

82. Brief of Amici Curiae Electronic Frontier Foundation et al. in Support of Affirmance, *In re Application of the U.S. for an Order Directing a Provider of Elec. Comm’n Serv. to Disclose Records to the Gov’t*, 620 F.3d 304 (3d Cir. 2010) (No. 08-4227), 2009 WL 3866619 (arguing that “when a cell phone user makes a call, the only information that is voluntarily and knowingly conveyed to the phone company is the number that is dialed”). *But see In re U.S. for Historical Cell Site Data*, 724 F.3d 600, 612 (5th Cir. 2013) (finding that “users know that they convey information about their location to their service providers when they make a call and that they voluntarily continue to make such calls”).

83. *Smith*, 442 U.S. at 743 (quoting *Katz v. United States*, 389 U.S. 347, 361 (1967)).

84. *See, e.g.*, 18 U.S.C. § 2703(a)–(c) (2012) (differentiating the standards for required disclosure of customer communications, 18 U.S.C. § 2703(a)–(b), versus records, 18 U.S.C. § 2703(c)); *Smith*, 442 U.S. at 741 (noting that pen registers differ significantly because they “do not acquire the *contents* of communications”) (emphasis added).

85. *A Guardian Guide to Your Metadata*, THE GUARDIAN (June 12, 2013, 11:52 AM), <http://www.theguardian.com/technology/interactive/2013/jun/12/what-is-metadata-nsa-surveillance#meta=1100000>.

86. *United States v. D’Andrea*, 497 F. Supp. 2d 117, 120 (D. Mass. 2007), *vacated*, 648 F.3d 1 (1st Cir. 2011); *see also Smith*, 442 U.S. at 741 (noting that pen registers differ significantly because they “do not acquire the *contents* of communication”); *In re U.S. for*

dialed and a sender's name, email address, and IP address). Thus, service providers can be forced to disclose this information to the government without implicating the Fourth Amendment. Conversely, the *contents* of communications, such as the actual message and items that are not directed to the third-party intermediary (such as the ISP), but rather to a specific recipient, are generally protected under the Fourth Amendment.⁸⁷ In *Katz* and its line of cases, where the service provider is merely the intermediary, the content-based communications between parties continue to give the user a reasonable expectation of privacy, and therefore the Fourth Amendment protects messages.⁸⁸

Taking into account the different roles of third parties in the content versus non-content scenarios, the Fourth Amendment analysis may just hinge on the position of the third party in relation to the information; specifically, “whether the third party created the record to memorialize its business transaction with the target” (non-content records), or rather the third party “simply record[ed] its observation of a transaction between two independent parties” (content data).⁸⁹ Where the third party simply records its observation of the transaction (such as recording email exchanges between parties), and is not a party to the transaction, the content is subject to Fourth Amendment protection.⁹⁰ However, where the third party collects and stores non-content information (such as recording customer subscription information like email to/from addresses or the telephone numbers dialed) for its own business purposes or for routing the communication, the third party becomes a party to the transaction and the Fourth Amendment does not protect such information. Thus, “[c]ommunications content, such as the contents of letters, phone calls, and emails, which are not directed to a business, but simply sent via that business, are generally protected. However, addressing information, which the business needs to route those communications appropriately and efficiently are not.”⁹¹

Historical Cell Site Data, 724 F.3d at 611–12 (holding that cell site information is a business record to which the cell service provider is a party and as such falls into the non-content bucket); *Guest v. Leis*, 255 F.3d 325, 336 (6th Cir. 2001) (holding that a user loses any expectation of privacy in personal subscription information when it is conveyed to a system operator).

87. *See In re U.S. for Historical Cell Site Data*, 724 F.3d at 611 (“Communications content, such as the contents of letters, phone calls, and emails, which are not directed to a business, but simply sent via that business, are generally protected. However, addressing information, which the business needs to route those communications appropriately and efficiently are not.”).

88. *See, e.g., Katz*, 389 U.S. at 352 (1967); *Warshak*, 631 F.3d at 286, 288.

89. *In re U.S. for Historical Cell Site Data*, 724 F.3d at 611.

90. *See, e.g., Warshak*, 631 F.3d at 286, 288.

91. *In re U.S. for Historical Cell Site Data*, 724 F.3d at 611.

3. *The Employer Context*

Expectations of privacy become more complicated when the third-party intermediary is an employer. Even more uncertainty abounds if the employer has a policy of monitoring electronic communications. This is because the reasonableness of a privacy expectation in the employment setting “is understood to differ according to context.”⁹² Nevertheless, individuals do not lose all Fourth Amendment rights in the employment setting.

“In both the public and private sectors, employees’ privacy rights are governed by whether or not the employee had a reasonable expectation of privacy.”⁹³ Although private sector employers are bound by the common law right to privacy—most often, a tort for the invasion of privacy—and as such fall outside of the scope of this Article, in the public sector, employer privacy rights are defined by constitutional principles under the Fourth Amendment.⁹⁴ This distinction is important because there are “more than 23 million government (federal, state, and local) employees in the United States, and, at least as a default matter, they have some claim to privacy while at work.”⁹⁵

*O’Connor v. Ortega*⁹⁶ sets forth the analytical framework for Fourth Amendment claims against public sector employers. There, the Supreme Court noted the importance of balancing the legitimate privacy interests of public employees with the realities of the workplace.⁹⁷ Ultimately, the Court

92. *O’Connor v. Ortega*, 480 U.S. 709, 715 (1987).

93. Molly DiBianca, *Is There a Reasonable Expectation of Privacy in Your Tweets?*, DEL. EMP. L. BLOG (July 23, 2013), <http://www.delawareemploymentlawblog.com/2013/07/is-there-a-reasonable-expectation-of-privacy-in-your-tweets.html>; *see also* TIMOTHY P. GLYNN, RACHEL ARNOW-RICHMAN, AND CHARLES A. SULLIVAN, *EMPLOYMENT LAW: PRIVATE ORDERING AND ITS LIMITATIONS* 349–50 (2d ed. 2011) (“In order for an employee or applicant to have a cognizable breach of privacy claim (of any kind), he or she must have a legitimate or reasonable expectation of privacy in the sphere upon which the employer intruded.”). However, “the sources of privacy protection analyzed in each case [(i.e., public employer versus private employer)] are different.” *Id.* Many courts faced with private sector privacy cases have borrowed from government employer cases, asking whether the employee had a reasonable expectation of privacy. *See, e.g.*, *O’Bryan v. KTIV Television*, 868 F. Supp. 1146, 1159 (N.D. Iowa 1994); *K-Mart Corp. Store No. 7441 v. Trotti*, 677 S.W.2d 632, 638 (Tex. App. 1984).

94. *See, e.g.*, *O’Connor*, 480 U.S. at 717 (“Individuals do not lose Fourth Amendment rights merely because they work for the government instead of a private employer.”); *Nat’l Treasury Emp. Union v. Von Raab*, 489 U.S. 656, 665 (1989) (“Our earlier cases have settled that the Fourth Amendment protects individuals from unreasonable searches conducted by the Government, even when the Government acts as an employer.”); DiBianca, *supra* note 93.

95. TIMOTHY P. GLYNN, RACHEL ARNOW-RICHMAN, AND CHARLES A. SULLIVAN, *EMPLOYMENT LAW: PRIVATE ORDERING AND ITS LIMITATIONS* 301 (2d ed. 2011).

96. 480 U.S. 709 (1987).

97. *Id.* at 721.

found that a physician at a state hospital did not lose his Fourth Amendment rights solely because of his decision to work for the government.⁹⁸ Instead, the physician enjoyed an expectation of privacy in his desk and file cabinets, as he did not share these items with other employees.⁹⁹ The Court, however, disagreed on precisely how to apply the reasonable expectation of privacy analysis in the public sector employment context, and the test remains unclear to this day.¹⁰⁰ In fact, in the most recent Supreme Court case on this issue, the Court proceeded with care in determining privacy expectations for employer-owned electronic equipment.¹⁰¹

More often than not, employer policies, or lack thereof, will shape the reasonableness of a privacy expectation.¹⁰² In many cases where public employers do not have formal policies allowing inspections of employee activities, courts have found both a subjective and objective reasonable expectation of privacy,¹⁰³ though not always a violation of the Fourth Amendment.¹⁰⁴ However, as a condition of employment and through the use of employer-provided technology, many employees effectively waive their

98. *Id.* at 718–19.

99. *Id.*

100. *See id.*; *see, e.g.*, *City of Ontario v. Quon*, 560 U.S. 746, 759 (2010) (finding that it was preferable to dispose of the case on narrower grounds because, “[e]ven if the Court were certain that the *O’Connor* plurality’s approach were the right one, the Court would have difficulty predicting how employees’ privacy expectations will be shaped” by rapidly changing technology).

101. *Id.* at 759 (“The judiciary risks error by elaborating too fully on the Fourth Amendment implications of emerging technology before its role in society has become clear.”).

102. *See, e.g., id.* at 760 (finding that “employer policies concerning communications will of course shape the reasonable expectations of their employees, especially to the extent that such policies are clearly communicated”).

103. *See, e.g., Schowengerdt v. Gen. Dynamics Corp.*, 823 F.2d 1328, 1335 (9th Cir. 1987) (finding that an employee “would enjoy a reasonable expectation of privacy in areas given over to his exclusive use, unless he was on notice from his employer that searches of the type to which he was subjected might occur from time to time for work-related purposes”), *abrogated by* *Pollard v. The GEO Grp., Inc.*, 629 F.3d 843 (9th Cir. 2012), *rev’d sub nom.* *Minneci v. Pollard*, 132 S. Ct. 617 (2012); *Gillard v. Schmidt*, 579 F.2d 825, 828 (3d Cir. 1978) (holding that “in the absence of an accepted practice or regulation to the contrary,” a school employee “[w]orking in an office secured by a locked door” had a reasonable expectation of privacy in his desk and its contents).

104. This is largely due to the plurality’s opinion in *O’Connor* stating that internal investigations into work-related misconduct should instead be judged by a standard of reasonableness. *See O’Connor*, 480 U.S. at 725–26; *United States v. Taketa*, 923 F.2d 665, 673–74 (9th Cir. 1991) (finding the search of an employee’s office subject to Fourth Amendment restraints but where the search was directed at uncovering work-related misconduct, the warrantless search was in-line with the *O’Connor* Court’s analysis and not in violation of the Fourth Amendment).

expectation of privacy and are effectively on notice of employer intrusion into their private space (both physical and electronic).¹⁰⁵

In cases where public employers have formal policies that sufficiently provide notice to employees that their communications and files may be monitored, employees lose their objectively reasonable expectation of privacy.¹⁰⁶ *United States v. Simmons* is a perfect example.¹⁰⁷ There, the government employer had an internet policy that clearly stated that the employer would “audit, inspect, and/or monitor an employee’s use of the internet.”¹⁰⁸ As a result, the court concluded that such “office practices, procedures, or regulations may reduce legitimate privacy expectations.”¹⁰⁹ Regardless of whether the employee had a “subjective” expectation of privacy in his internet usage, the internet policy effectively negated an “objective” expectation of privacy.¹¹⁰ Subsequently, when the government employer conducted remote searches of the employee’s computer, the employee’s Fourth Amendment rights were not violated, since the employee lacked “a legitimate expectation of privacy in the files downloaded from the internet.”¹¹¹

4. *What about Passwords?*

The status of passwords within the morass of the third-party doctrine remains unknown. Under the content versus non-content distinction noted above in Subsection III.B.2, one might argue that passwords should be

105. See, e.g., *O’Connor*, 480 U.S. at 717 (“Public employees’ expectations of privacy . . . may be reduced by virtue of actual office practices and procedures.”); Pauline T. Kim, *Electronic Privacy and Employee Speech*, 87 CHI.-KENT L. REV. 901, 919 (2012); Lewis Maltby, *Employment Privacy: Is There Anything Left*, HUM. RTS., 2013, available at http://www.americanbar.org/publications/human_rights_magazine_home/2013_vol_39/may_2013_n2_privacy/employment_privacy.html (“The actual test of whether an employee has a reasonable expectation of privacy is who owns the equipment used to transmit the message. If the equipment belongs to the employer, the employer has the right to monitor anything and everything on it.”).

106. See, e.g., *Biby v. Bd. of Regents of Univ. of Neb. at Lincoln*, 419 F.3d 845, 850–51 (8th Cir. 2005) (finding an employee does not have a reasonable expectation of privacy in his computer files where the employer’s computer policy allows for searches); *United States v. Angevine*, 281 F.3d 1130, 1134 (10th Cir. 2002) (finding that the employer’s “policies and procedures prevent its employees from reasonably expecting privacy in data downloaded from the Internet onto University computers”); *United States v. Hamilton*, 778 F. Supp. 2d 651, 654 (E.D. Va. 2011) (finding that a public school employee lacked an objectively reasonable expectation of privacy in emails stored on his work computer since he had been on notice that contents of his computer were subject to inspection).

107. 206 F.3d 392, 398–99 (4th Cir. 2000).

108. *Id.* at 398.

109. *Id.*

110. *Id.*

111. *Id.*

considered non-content information and therefore their compelled disclosure would not violate an individual's Fourth Amendment rights. This is a plausible argument; a password for an electronic device resembles telephone digits entered into a phone, as passwords often have to be entered in order to actually utilize the device. And, passwords are, after all, the key to unlocking the content, and a third party can theoretically record them. However, when viewed in totality, passwords should be treated like the *contents* of communication. Although passwords in the traditional sense may have functioned simply as a key to unlocking data, such an analogy fails in the digital age. At least in the form of decryption keys, passwords change content from unreadable to readable text, thereby communicating information.¹¹² Despite the uncertainty of whether passwords are content or non-content data, the role of third parties in relation to passwords is not "sufficient to extinguish a reasonable expectation of privacy,"¹¹³ and thus should help to correctly place passwords in the *Katz* line of cases.

First, numerous third-party service providers do not record password information. Instead, third-party service providers often have only the ability to reset the password should a user forget it.¹¹⁴ Second, although third-party employers and some service providers may actually have access to this information,¹¹⁵ individuals do not assume that passwords are used by their

112. See *In re Under Seal*, 749 F.3d 276, 279 (4th Cir. 2014) (noting that decryption is the process of changing ciphertext (an unreadable jumble of letters and numbers) back into plaintext (readable data)); see also Brief for the ACLU Found. of Mass et al. as Amici Curiae Supporting Defendant at 5–7, 19–20, *Commonwealth v. Gelfgatt*, 11 N.E.3d 605 (Mass. 2014) (No. SJC-11358), available at https://www.eff.org/files/2013/10/29/brief_of_amici_curiae_aclu_aclu_eff.pdf.

113. *United States v. Warshak*, 631 F.3d 266, 286 (6th Cir. 2010).

114. See, e.g., JOSEPH BONNEAU & SÖREN PREIBUSCH, *THE PASSWORD THICKET: TECHNICAL AND MARKET FAILURES IN HUMAN AUTHENTICATION ON THE WEB* 19 (2010), available at http://weis2010.econinfosec.org/papers/session3/weis2010_bonneau.pdf ("The best solution, sending a time-limited reset link . . . was implemented about half of the time, with identity sites being very significantly more likely to implement this."); *Apple ID: Changing Your Password*, APPLE SUPPORT, <http://support.apple.com/en-us/HT201355> (last modified Dec. 11, 2014) (detailing steps to changing your Apple ID password) [hereinafter *Apple ID*]; *Customer Proprietary Network Information (CPNI) for Wireless Consumers*, VERIZON, <http://www.verizonwireless.com/b2c/globalText?contentType=Legal%20Notice&textId=181> (last visited Jan. 3, 2015) (noting that the information collected includes "services purchased (including specific calls you make and receive), related local and toll billing information, the type, destination, technical configuration, location and amount of use of purchased services," but does not mention passwords) [hereinafter *CPNI*]. But see *AT&T Privacy Policy*, AT&T, <http://www.att.com/gen/privacy-policy?pid=2506> (last updated Sept. 16, 2013) (noting that AT&T collects account information such as security codes).

115. With regard to the employment context, see *Keylogging Employees' Computer Use Met with Judicial Wariness*, JACKSON LEWIS (June 5, 2009), <http://www.jacksonlewis.com/legalupdates/article.cfm?aid=1747> (discussing a case in which an employer used keylogging to discover an employee's password) [hereinafter JACKSON LEWIS].

providers for any legitimate business purpose, nor do users of electronic devices necessarily know that third parties have facilities for making permanent records of the passwords they enter.¹¹⁶ Additionally, users of electronic devices do not voluntarily hand over their password when entering it into their device.¹¹⁷ Instead, users assume that only they know the password,¹¹⁸ since “[t]he important feature about PINs and passwords is that they’re generally something that we know”¹¹⁹ and that have to be recalled through personal memory and usually cannot be directly regained through a “forgot my password” option. Finally, although less clear, third parties may be viewed as mere intermediaries because third parties do not create password records to memorialize the business transaction with the individual; the user creates his own password independent of the business transaction with his service provider.¹²⁰

Moreover, because most service providers do not even collect or store passwords,¹²¹ the third-party doctrine should not be considered in assessing a user’s reasonable expectation of privacy in his passwords. Even if some service providers collect passwords, customers generally are not aware that phone companies are recording this information,¹²² thereby negating the

116. Brief of Amici Curiae Electronic Frontier Foundation et al., *In re Application of U.S. for an Order Directing a Provider of Elec. Comm’n Serv. to Disclose Records to the Gov’t*, 620 F.3d 304 (3d Cir. 2010) (No. 08-4227), 2009 WL 386619 (arguing that “when a cell phone user makes a call, the only information that is voluntarily and knowingly conveyed to the phone company is the number that is dialed”). However, it is worth noting that “[t]o the extent an individual knowingly exposes his activities to third parties, he surrenders Fourth Amendment protections.” Reporters Comm. for Freedom of the Press v. Am. Tel. & Tel. Co., 593 F.2d 1030, 1043 (D.C. Cir. 1978) (emphasis added).

117. See *supra* Section III.B.

118. See, e.g., CPNI, *supra* note 114; *Apple ID*, *supra* note 114. But see *AT&T Privacy Policy*, *supra* note 114.

119. Marcia Hoffmann, *Apple’s Fingerprint ID May Mean You Can’t “Take the Fifth,”* WIRED (Sept. 12, 2013), <http://www.wired.com/opinion/2013/09/the-unexpected-result-of-fingerprint-authentication-that-you-cant-take-the-fifth/>.

120. *In re U.S. for Historical Cell Site Data*, 724 F.3d 600, 611 (5th Cir. 2013) (noting that “the right to possession hinges on whether the third party created the record to memorialize its business transaction with the target”). For a discussion of the implications of third parties who serve as intermediaries, see *supra* Subsection III.B.2.

121. Angwin, *supra* note 2, at B4 (“Spokeswomen for Microsoft Corp. and Research In Motion Ltd. say their companies don’t collect or store passwords.”). Even if service providers “store” passwords, in technical terms, they are “hashed and salted”—a one-way cryptographic strategy that allows an application to authenticate a user without the ability to read their password. *Salted Password Hashing—Doing it Right*, CRACKSTATION, <https://crackstation.net/hashing-security.htm> (last modified Aug. 6, 2014, 7:12 PM).

122. Brief of Amici Curiae Electronic Frontier Foundation et al., *In re Application of U.S. for an Order Directing a Provider of Elec. Comm’n Serv. to Disclose Records to the Gov’t*, 620 F.3d 304 (3d Cir. 2010) (No. 08-4227), 2009 WL 386619 (arguing that “when a cell phone user makes a call, the only information that is voluntarily and knowingly conveyed

third-party doctrine. Courts should therefore treat passwords like email given the similar nature of the sensitive information conveyed in both, and should determine that the “the degree of access granted to [third parties] does not diminish the reasonableness”¹²³ of privacy in passwords.

Despite the argument that passwords should be protected under the Fourth Amendment, there might be a caveat in the employment context. Because an employee’s expectation of privacy in an employment setting depends heavily on an employer’s policies, practices, procedures and regulations,¹²⁴ whether employees have an expectation of privacy in their passwords on employer-owned electronic devices is subject to a case-by-case analysis. When an employer has a policy notifying employees that both their electronic communications will be monitored and that keystroke logging will be used to track and record *anything* typed into a device, courts should assume that such employees do not have an “objectively” reasonable expectation of privacy. In these limited circumstances, courts should employ an employment caveat for privacy protections and not afford Fourth Amendment protection to passwords. But in a situation where an employer’s policies and procedures give employees notice that their computer-activity is being monitored but do not mention the use of keystroke logging (an important practice in being able to record passwords typed into a device), a grey area would remain. In this grey area, an employee should still have a reasonable expectation of privacy, both subjective and objective, in passwords typed into employer-provided devices, until keystroke logging is common practice and the subject of widespread awareness in all workplace settings,¹²⁵ thereby negating an “objectively” reasonable expectation.

to the phone company is the number that is dialed”). *But see In re U.S. for Historical Cell Site Data*, 724 F.3d at 612 (finding that “users know that they convey information about their location to their service providers when they make a call and that they voluntarily continue to make such calls”).

123. *See* *United States v. Warshak*, 631 F.3d 266, 287 (6th Cir. 2010).

124. *See, e.g., O’Connor v. Ortega*, 480 U.S. 709, 717 (1987) (“Public employees’ expectations of privacy . . . may be reduced by virtue of actual office practices and procedures, or by legitimate regulation.”); *United States v. Angevine*, 281 F.3d 1130, 1134–35 (10th Cir. 2002) (citing *O’Connor v. Ortega* and noting that “Oklahoma State University policies and procedures prevent its employees from reasonably expecting privacy in data downloaded from the Internet onto University computers”); *People v. Kent*, 910 N.Y.S.2d 78, 92–93 (2010) (“An employee’s expectation of privacy in material stored in an office computer depends upon the employer’s policy regarding computer use and any other relevant office practices, procedures, and regulations.”).

125. As it is, keystroke logging “is done secretly, so the person using the keyboard is unaware his activities are being monitored.” JACKSON LEWIS, *supra* note 115.

IV. ATTEMPTS AT PROTECTING PASSWORDS UNDER THE FIFTH AMENDMENT

Although the Fourth Amendment protects individuals by keeping secure their persons, papers, and effects against unreasonable seizures, it still allows the government to obtain Fourth Amendment-protected information through a narrowly-tailored search warrant based on probable cause.¹²⁶ But if the information compelled by the search warrant is testimonial and incriminating, then the legal analysis shifts to the Fifth Amendment.

A. ESTABLISHING THE REQUIREMENTS TO INVOKE THE FIFTH AMENDMENT

Under the Fifth Amendment, “[n]o person shall be . . . compelled in any criminal case to be a witness against himself.”¹²⁷ In order to invoke this right, an individual must establish three things: (1) compulsion, (2) a testimonial communication, and (3) incrimination.¹²⁸

(1) Compulsion. Under the Fifth Amendment, compulsion requires that the government force an individual to surrender information.¹²⁹ Essentially, courts must decide if testimony is free and voluntary or if it was obtained through improper influence.¹³⁰ If the latter is found to be true, courts view this “extortion of information from the accused himself [as] offen[sive] [to] our sense of justice.”¹³¹

(2) Testimonial. In order for the act to be “testimonial” under the Fifth Amendment, the result of revealing the information must cause a person to be a “witness” against himself.¹³² Essentially, the communication must force the individual into “the cruel trilemma,” compelling the individual to choose between self-accusation, perjury or contempt when asked to give a sworn

126. Kerr, *supra* note 76, at 1211–14.

127. U.S. CONST. amend. V.

128. Shields, *supra* note 11.

129. See, e.g., Fisher v. United States, 425 U.S. 391, 397 (1976) (finding the Fifth Amendment serves to prohibit the use of “physical or moral compulsion” on a person asserting privilege); Boyd v. United States, 116 U.S. 616, 630 (1886) (holding that “any forcible and compulsory extortion of a man’s own testimony, or of his private papers to be used as evidence to convict him of crime” violated the Fourth and Fifth Amendments).

130. See Bram v. United States, 168 U.S. 532, 542–43 (1897) (holding that the constitutional inquiry is whether the confession was “free and voluntary: that is, must not be extracted by any sort of threats or violence, nor obtained by any direct or implied promises, however slight, nor by the exertion of any improper influence”) (internal citations omitted).

131. Couch v. United States, 409 U.S. 322, 328 (1973).

132. See Doe v. United States, 487 U.S. 201, 210 (1988) (“[I]n order to be testimonial, an accused’s communication must itself, explicitly or implicitly, relate a factual assertion or disclose information. Only then is a person compelled to be a ‘witness’ against himself.”).

statement containing factual information.¹³³ This privilege does not extend to the production of real or physical evidence,¹³⁴ but does protect expressions of content in an individual's mind.¹³⁵ And an act is not testimonial if the information revealed is already known by the government—a “foregone conclusion.”¹³⁶

(3) Incrimination. Additionally, the compelled testimonial communication must be incriminating. Courts must determine whether the revelation of the testimony would create a substantial hazard of self-incrimination or otherwise expose an individual to a criminal charge.¹³⁷

However, an important limitation of the Fifth Amendment must be noted: the protection of the Fifth Amendment is limited to when an individual is “compelled to be a witness against himself.”¹³⁸ The Supreme Court has repeatedly held that the Fifth Amendment was never intended to allow a person to argue that some other person “might be incriminated by his testimony.”¹³⁹ This is true even if that person can be classified as an agent of the other person.¹⁴⁰ Thus, the incrimination provision may allow individuals to claim Fifth Amendment protection, but it leaves no redress for third parties compelled to hand over passwords to the government.

B. COMPELLING PASSWORDS FROM INDIVIDUALS: DOESN'T THAT VIOLATE THE FIFTH AMENDMENT?

Lately, the government has become aggressive in its quest for passwords.¹⁴¹ And while individuals have been quick to fight back with the Fifth Amendment privilege, courts have diverged in deciding whether the privilege does in fact protect passwords.

In 2012 the Eleventh Circuit ruled on the issue of whether a defendant could be required to decrypt an encrypted hard drive without implicating the Fifth Amendment.¹⁴² In that case, the defendant was served with a subpoena

133. *See* *Pennsylvania v. Muniz*, 496 U.S. 582, 596 (1990).

134. *See id.* at 589 (“[W]e have long held that the privilege does not protect a suspect from being compelled by the State to produce ‘real or physical evidence.’”).

135. *See Doe*, 487 U.S. at 210 n.9.

136. *United States v. Hubbell*, 530 U.S. 27, 44 (2000).

137. *See United States v. Reis*, 765 F.2d 1094, 1095 (11th Cir. 1985) (the witness “must be faced with substantial and real hazards of self-incrimination”); *Hale v. Henkel*, 201 U.S. 43, 67 (1906).

138. *Fisher v. United States*, 425 U.S. 391, 398 (1976).

139. *Hale*, 201 U.S. at 69; *see also Fisher*, 425 U.S. at 398.

140. *See Fisher*, 425 U.S. at 397.

141. *See, e.g., Commonwealth v. Gelfgatt*, 11 N.E.3d 605, 614 (Mass. 2014); Angwin, *supra* note 2, at B4; Farivar, *supra* note 2.

142. *In re Grand Jury Subpoena Duces Tecum Dated Mar. 25, 2011*, 670 F.3d 1335 (11th Cir. 2012).

because forensic examiners were unable to view the encrypted portions of his seized laptops and external hard drives.¹⁴³ The subpoena required him to produce the unencrypted *contents* located on the hard drives of his laptops and five external hard drives.¹⁴⁴ He refused to comply, invoking his Fifth Amendment privilege against self-incrimination.¹⁴⁵ Although the government conceded that the decryption and production of the data would be “compelled” and “incriminatory,” the government disputed whether it would also be “testimonial.”¹⁴⁶ The Eleventh Circuit, reversing the district court, found that the defendant’s “decryption and production of the hard drives’ contents would trigger Fifth Amendment protection because it would be testimonial.”¹⁴⁷ Despite the government’s argument that it was not seeking “the combination or key, but rather the contents,” the court found that production would still require the individual to use “the contents of his own mind” to decrypt the files.¹⁴⁸ The court centered its analysis on the act of production, finding the Fifth Amendment protected against production and decryption of the hard drives.¹⁴⁹

Notwithstanding the Eleventh Circuit’s guidance, a state supreme court has recently ruled that individuals can be compelled to divulge passwords when the information that would be disclosed is a “foregone conclusion.”¹⁵⁰ In *Commonwealth v. Gelfgatt*, the Massachusetts Supreme Court conceded that entering an encryption key would appear to be a “testimonial communication. . . trigger[ing] Fifth Amendment protection,” but noted that it also had to determine whether the act of production would “lose[] its testimonial character” because the information would be a “foregone conclusion.”¹⁵¹ And because the defendant’s “ownership and control of the computers,” “knowledge of the fact of encryption,” and “knowledge of the encryption key” were already known, the court concluded that entering an encryption key in this case was a “foregone conclusion” and therefore was not a testimonial communication protected by the Fifth Amendment.¹⁵²

143. *Id.* at 1340.

144. *Id.*

145. *Id.* at 1337.

146. *Id.* at 1341–42.

147. *Id.* at 1341.

148. *Id.* at 1345–46. The Supreme Court has long held that forcing the accused “to disclose the contents of his own mind” implicates the Fifth Amendment. *See, e.g.,* *Curcio v. United States*, 354 U.S. 118, 128 (1957).

149. *In re Grand Jury Subpoena Duces Tecum Dated Mar. 25, 2011*, 670 F.3d at 1346.

150. *Commonwealth v. Gelfgatt*, 11 N.E.3d 605, 614 (Mass. 2014).

151. *Id.*

152. *Id.* at 615.

This split in opinion has played out in the district courts over the last several years. While some courts have agreed with the Eleventh Circuit, finding that compelling an individual's password would infringe on his Fifth Amendment self-incrimination privilege,¹⁵³ others have found exceptions to an individual's claim of Fifth Amendment privilege. In the former category lies *United States v. Kirschner*, where the government subpoenaed a defendant to divulge the password to his computer,¹⁵⁴ but the court found that providing the password would be like giving the combination to a safe and would be a testimonial communication.¹⁵⁵ In contrast, other courts have bypassed the Fifth Amendment privilege by: (1) compelling defendants to provide an unencrypted version of the hard drive, but not the passwords, where the subpoena requested passwords to the computer;¹⁵⁶ (2) being amenable to government requests for the *contents* of the encrypted data;¹⁵⁷ and (3) finding that decryption is not testimonial when the information sought to be obtained is already known.¹⁵⁸ In these cases, only the underlying data, not the passwords, were revealed.¹⁵⁹ However, as the approaches taken by the lower courts further diverge, the issue may soon make its way up to the Supreme Court.

153. *United States v. Kirschner*, 823 F. Supp. 2d 665, 668 (E.D. Mich. 2010); *see also* *United States v. Rogozin*, No. 09-CR-379(S)(M), 2010 WL 4628520, at *6 (W.D.N.Y. Nov. 16, 2010) (holding that the defendant's statement as to his password was testimonial in nature and recommending suppression of the statement).

154. *Kirschner*, 823 F. Supp. 2d at 668.

155. *Id.*

156. *See, e.g., In re Boucher (Boucher II)*, No. 2:06-mj-91, 2009 WL 424718, at *1, *4 (D. Vt. Feb. 19, 2009) (directing the defendant to provide an unencrypted version of hard drive where the government's expert was unable to search a computer because of password-protection and where the subpoena had asked for "any passwords associated with the laptop"). It is worth noting that in *In re Boucher (Boucher I)*, No. 2:06-mj-91, 2007 WL 4246473 (D. Vt. Nov. 29, 2007), the magistrate judge found the act of producing the password was testimonial and privileged, and as a result, on appeal, the request for a password was revised to a request for an unencrypted version of the hard drive. *Boucher II*, 2009 WL 424718, at *2.

157. *See* *United States v. Fricosu*, 841 F. Supp. 2d 1232, 1235, 1238 (D. Colo. 2012) (holding that defendant shall provide an unencrypted copy of her computer's hard drive to the government where the government sought a writ for the unencrypted contents).

158. *See, e.g., In re Decryption of a Seized Data Storage System*, No. 13-M-449, at 3 (E.D. Wis. filed May 21, 2013) (order granting ex parte request for reconsideration), *available at* http://www.wired.com/images_blogs/threatlevel/2013/05/decryptorder.pdf.

159. But, at least one court has forced an individual to hand over their passwords to social networking sites. *See* *McMillen v. Hummingbird Speedway Inc.*, No. 113-2010 CD, 2010 Pa. Dist. & Cnty. Dec. LEXIS 270, at *13 (Pa. Ct. Com. Pl. Sept. 9, 2010) (ordering plaintiff to "provide his Facebook and MySpace user names and passwords to counsel").

Finally, it is important to recognize that a password or an encryption key alone would not by itself be incriminating—only testimonial.¹⁶⁰ However, when the information behind the “lock” is incriminating, providing the combination would trigger the privilege.¹⁶¹ And while password “locks” could be written down, they usually lie solely in one’s mind—thus causing an individual to use “the contents of his own mind” to divulge this information, and thereby triggering Fifth Amendment testimonial protections.¹⁶² However, if the government were to require manufacturers or users of encryption keys to keep a copy of the keys,¹⁶³ the testimonial compulsion would likely disappear along with the Fifth Amendment privilege, as the password would no longer lie solely in one’s mind. Likewise, if password authentication systems move from combinations to biometrics, such as fingerprints (a likely possibility given Apple’s use of fingerprints to unlock some iPhones), the government could demand biometric passwords without implicating the Fifth Amendment¹⁶⁴ because the Supreme Court has decided that biometrics are not testimonial.¹⁶⁵

160. See, e.g., *Fisher v. United States*, 425 U.S. 391, 410 (1976) (concluding that an act of production is testimonial if it concedes the existence, possession, control and authenticity of the documents); *In re Grand Jury Subpoena Duces Tecum Dated Mar. 25, 2011*, 670 F.3d 1335, 1345 (11th Cir. 2012) (holding that “an act of production can be testimonial when that act conveys some explicit or implicit statement of fact that certain materials exist, are in the subpoenaed individual’s possession or control, or are authentic”); Vivek Mohan & John Villasenor, *Decrypting the Fifth Amendment: The Limits of Self-Incrimination in the Digital Era*, 15 U. PA. J. CONST. L. HEIGHT. SCRUTINY 11, 24 (2012) (“Courts generally agree that divulging a password constitutes a testimonial act.”).

161. Engel, *supra* note 11, at 556.

162. Hoffmann, *supra* note 119, at 2 (“These memory-based authenticators are the type of fact that benefit from strong Fifth Amendment protection should the government try to make us turn them over against our will.”).

163. *Privacy in the Digital Age: Encryption and Mandatory Access: Hearing Before the Subcomm. on the Constitution, Federalism, & Prop. Rights of the Comm. on the Judiciary*, 105th Cong. 14 (1998) (testimony of Robert S. Litt, Principal Assoc. Deputy Att’y Gen.), available at <http://www.loc.gov/law/find/hearings/pdf/00139296461.pdf>.

164. See, e.g., *Virginia v. Baust*, No. CR14-1439, 2014 WL 6709960, at *3 (Va. Cir. Ct. Oct. 28, 2014) (holding that a “[d]efendant cannot be compelled to produce his passcode to access his smartphone but he can be compelled to produce his fingerprint to do the same,” because his “fingerprint, like a key, . . . does not require the witness to divulge anything through his mental processes”); Hoffmann, *supra* note 119, at 2 (“[I]f we move toward authentication systems based solely on physical tokens or biometrics—things we have or things we are, rather than things we remember—the government could demand that we produce them without implicating anything we know.”); Jack Linshi, *Why the Constitution Can Protect Passwords But Not Fingerprint Scans*, TIME (Nov. 6, 2014), <http://time.com/3558936/fingerprint-password-fifth-amendment/>.

165. See *Gilbert v. California*, 388 U.S. 263, 266–67 (1967) (finding that a suspect may be compelled to provide a handwriting exemplar); *Schmerber v. California*, 384 U.S. 757, 765 (1966) (finding that a suspect may be compelled to furnish a blood sample).

C. ADDITIONAL CIRCUMVENTIONS OF THE FIFTH AMENDMENT

Even if the Supreme Court were to hear a case to resolve the divergent approaches of the courts and find a Fifth Amendment privilege for individuals, the question remains: what is left of an individual's Fifth Amendment password protection when it can be circumvented by getting the information from a third party? All of the cases above involve instances where the government compelled passwords from the creator of the password. They do not address how a court would analyze a situation in which a password is stored on the servers of an innocent third party. Because the Fifth Amendment is "not implicated when the government tries to compel passwords from third parties,"¹⁶⁶ the fact that the data is located on a third party's servers may destroy the individual's Fifth Amendment protections. The text of the Fifth Amendment specifically limits this privilege to one compelled to be a "witness against himself."¹⁶⁷ And the Supreme Court has found that the Fifth Amendment does not play any role when a third party is resisting disclosure of information to protect somebody else.¹⁶⁸ In essence, the Fifth Amendment provides a weak protection for individuals in protecting their digital footprint. The government's ability to compel passwords from third parties undermines the legal rights individuals currently enjoy under the Fifth Amendment and essentially renders the Fifth Amendment useless in protecting an individual's digital incriminating communications.

Although technology has changed, privacy rights should not. But as the law currently stands, passwords can be obtained by carefully navigating around two constitutional amendments that implicate privacy concerns. Even if passwords are protected under the Fourth Amendment, they can still be obtained through a narrowly-tailored search warrant based on probable cause. And, while an individual may then be able to invoke the Fifth Amendment to protect their passwords from disclosure, a third party could not. Thus, passwords could be easily obtained by serving a search warrant on a third party, complying with the Fourth Amendment without implicating the Fifth Amendment—a troublesome notion that the drafters of the Fifth Amendment surely did not anticipate. If the Fifth Amendment will not bar

166. Orin Kerr, *US Government Getting Password Information? (And Why the Story Raises More Questions than Answers)*, VOLOKH CONSPIRACY (July 26, 2013, 2:43 PM), <http://www.volokh.com/2013/07/26/u-s-government-getting-password-information-and-why-the-story-raises-more-questions-than-answers/>.

167. U.S. CONST. amend. V.

168. *See, e.g., Fisher v. United States*, 425 U.S. 391, 398 (1976); *Hale v. Henkel*, 201 U.S. 43, 69 (1906).

the government from the compulsory extortion of an individual's private passwords from a third party, then other laws should step in.

V. INADEQUACY OF THE ELECTRONIC COMMUNICATIONS PRIVACY ACT

As technology has evolved, the Electronic Communications Privacy Act (ECPA)¹⁶⁹ has not. Enacted in 1986¹⁷⁰—when the commercial internet did not exist and cellphones were too heavy and expensive to be easily used—the ECPA purports to protect the privacy of Americans' electronic lives.¹⁷¹ Congress enacted the ECPA “to update and clarify Federal privacy protections and standards in light of dramatic changes in new computer and telecommunication technologies,”¹⁷² intending to “fairly balance . . . the interests of privacy and law enforcement.”¹⁷³ Congress enacted the Stored Communications Act (“SCA”) as Title II of the ECPA to address voluntary and compelled disclosure of “stored wire and electronic communication and transactional records” held by third-party service providers because the internet created a host of privacy issues for users that may not have been adequately protected by the Fourth Amendment.¹⁷⁴ The SCA “creates rights [for] ‘customers’ and ‘subscribers’ of . . . service providers in both content and non-content information held by two particular types of providers:”¹⁷⁵ electronic communication service (“ECS”) ¹⁷⁶ providers (such as telephone

169. 18 U.S.C. §§ 2510–2522 (2012).

170. The ECPA was enacted just two years after Super Mario burst into homes on Nintendo (1984), *Nintendo History*, NINTENDO, <https://www.nintendo.co.uk/Corporate/Nintendo-History/Nintendo-History-625945.html> (last visited Jan. 6, 2015); three years before Zack Morris was spotted carrying a five-pound cellphone on Saved by The Bell (1989), ZACK MORRIS CELL PHONE.COM, <http://zackmorriscellphone.com/> (last visited Jan. 2, 2015); and more than twenty years before Apple released the first iPhone (2007), *7 Years of the iPhone: An Interactive Timeline*, TIME (June 27, 2014), <http://time.com/2934526/apple-iphone-timeline/>.

171. See *Modernizing the Electronic Communications Privacy Act*, AM. CIVIL LIBERTIES UNION, <http://www.aclu.org/technology-and-liberty/modernizing-electronic-communications-privacy-act-ecpa> (last visited Aug. 14, 2013).

172. S. REP. NO. 99-541, at 1 (1986).

173. *Id.* at 50.

174. 18 U.S.C. §§ 2701–2712 (2012).

175. Kerr, *supra* note 76, at 1211, 1213.

176. 18 U.S.C. § 2510 (2012) (defining an electronic communication service as “any service which provides to users thereof the ability to send or receive wire or electronic communications”).

companies) and remote computing service (“RCS”)¹⁷⁷ providers (such as YouTube).

While the approaches mentioned in Parts III and IV primarily focus on compelling or gaining access to passwords from an individual, the analysis changes when a third-party service provider is subpoenaed—courts must consider the implications of the SCA. The SCA prevents a “provider” of communications from voluntarily disclosing communications or customer information to entities and individuals¹⁷⁸ and also *requires disclosure* of information in certain instances.¹⁷⁹ The SCA specifically requires a “provider of electronic communication service or remote computing service”¹⁸⁰ to disclose to the government “record[s] or other information pertaining to a subscriber to or customer of such service,”¹⁸¹ when the governmental entity obtains a warrant or court order,¹⁸² and to disclose the contents of the communications when the government complies with a specific framework.¹⁸³

But the SCA does not explicitly mention disclosure of “passwords” or even define what constitutes a “record or other information pertaining to a subscriber or customer.”¹⁸⁴ Even if some service providers classify passwords as “account information” alongside traditional forms of customer information, passwords may fall into the non-content provision of the SCA—if they lie anywhere within the SCA.¹⁸⁵ Without clear direction, lawyers, law enforcement, and third parties are left to guess whether the SCA covers disclosures of passwords by remote computing service or electronic communication service providers. This is not surprising given that electronic passwords did not exist when the ECPA and the SCA were enacted. But if passwords were found to be “electronic communications” (which “means

177. 18 U.S.C. § 2711 (2012) (defining a remote computing service as any service that provides the public with “computer storage or processing services by means of an electronic communications system”).

178. Kerr, *supra* note 76, at 1223 (“Providers of ECS or RCS to the public ordinarily cannot disclose either content or noncontent information.”).

179. *Id.* at 1224 (“One of the most fundamental distinctions in the SCA is the distinction between voluntary disclosure regulated by § 2702 and compelled disclosure regulated by § 2703.”).

180. 18 U.S.C. § 2703(c)(1) (2012).

181. *Id.*

182. *Id.* § 2703(c)(1)(A)–(B) (stating that for certain types of investigations and when the government is only seeking basic subscriber information not subject to the scope of this Article, a governmental entity need only submit a formal written request).

183. *Id.* § 2703(a)–(b).

184. *See id.* § 2703(c)(1).

185. Contrast this to “contents,” which “when used with respect to any wire, oral, or electronic communication, includes any information concerning the substance, purport, or meaning of that communication.” *Id.* § 2510(8).

any transfer of signs, signals, writing, images, sounds, data, or intelligence of any nature transmitted in whole or in part by a wire, radio, electromagnetic, photoelectronic or photooptical system that affects interstate or foreign commerce”),¹⁸⁶ governmental entities could still obtain this information from third-party service providers with a warrant.¹⁸⁷

Legal scholars should question the efficacy of the SCA in conjunction with the Fifth Amendment in the digital age and consider whether incriminating information should be further protected beyond the confines of the Fifth Amendment. If courts consider electronic passwords to be within the scope of the SCA, then governmental entities could use the SCA to compel disclosure of passwords from third-party providers (assuming that a court would find a service provider such as Google or Apple to be an ECS (“any service which provides to users thereof the ability to send or receive wire or electronic communications”¹⁸⁸)). If courts interpret the SCA this way, law enforcement could circumvent the Fifth Amendment by simply seeking information from third parties rather than the suspect.

VI. COMPELLING PASSWORDS FROM THIRD PARTIES

To date, only a few courts have ruled on whether individuals (or entities) can be forced to turn over passwords to their computer files.¹⁸⁹ Other courts

186. *Id.* § 2510(12).

187. *See id.* § 2703(a)–(b).

188. *Id.* § 2510(15). For example, “telephone companies and electronic mail companies” generally act as providers of electronic communication services. S. REP. NO. 99-541, at 14 (1986), *reprinted in* 1986 U.S.C.C.A.N. 3555, 3568.

189. In light of the rulings, only a few courts have specifically determined whether an individual is forced to turn over his password when law enforcement requests it. *See* United States v. Kirschner, 823 F. Supp. 2d 665, 668 (E.D. Mich. 2010) (compelling an individual’s password would infringe on his Fifth Amendment incrimination privilege); Commonwealth v. Gelfgatt, 11 N.E.3d 605, 614–15 (Mass. 2014) (holding that “entering an encryption key . . . would appear . . . to be a testimonial communication that triggers Fifth Amendment protection,” but that based on the facts in the case, entering an encryption key was a foregone conclusion, and as such, “the act of decryption is not a testimonial communication that is protected by the Fifth Amendment”); Virginia v. Baust, No. CR14-1439, 2014 WL 6709960, at *3 (Va. Cir. Ct. 2014) (finding that “compelling Defendant to provide access through his passcode is both compelled and testimonial and therefore protected”). Other courts have skirted the Fifth Amendment protection and allowed disclosure of information without compelling disclosure of the actual password. *See In re Under Seal*, 749 F.3d 276, 293 (4th Cir. 2014) (finding no “basis upon which to challenge the Pen/Trap order” where petitioner failed to properly raise an argument about the district court’s compulsion of encryption keys or the issue of plain error review); United States v. Fricosu, 841 F. Supp. 2d 1232, 1235, 1238 (D. Colo. 2012) (holding that defendant shall provide an unencrypted copy of her computer’s hard drive to the government where the government sought a writ for the unencrypted contents); *In re Boucher (Boucher II)*, No. 2:06-mj-91, 2009 WL 424718, at *1, *4 (D. Vt. Feb. 19, 2009) (directing the defendant to provide an unencrypted version of his

have wrangled with the question of whether individuals can be forced to turn over passwords to their social networking sites,¹⁹⁰ or whether individuals can be forced to unencrypt and produce the contents of their hard drives,¹⁹¹ but the issue as to whether governmental entities can compel service providers to hand over users' passwords remains unresolved. As mentioned in Part I, the government is already requesting users' passwords from service providers, and "there are novel and serious [F]ourth and [F]ifth Amendment issues raised by a policy that would compel the use of recoverable encryption"¹⁹² or some other form of third-party access to passwords.

A. SOLUTIONS FOR COMBATING COMPULSION OF PASSWORDS FROM SERVICE PROVIDERS

From the individual password holder's perspective, the government should not be able to get his password by subpoenaing a third party. But the Fourth and Fifth Amendments do not guarantee this prohibition and therefore do not adequately protect individuals when third parties are compelled to hand over passwords.

One possible solution to this problem is to require the government to demand an unencrypted copy of the protected data instead of the actual password.¹⁹³ Here, the government would only need to meet the search warrant requirements of the Fourth Amendment and use the SCA to obtain the information from a service provider, or subpoena the defendant to compel production of an unencrypted copy of the data (but only when it is a foregone conclusion that the incriminating evidence exists in the encrypted data).¹⁹⁴ However, this approach still has adverse privacy consequences. If

hard drive where the government's expert was unable to search a computer because of password-protection and where the subpoena had asked for "any passwords associated with the laptop").

190. *See, e.g.*, *McMillen v. Hummingbird Speedway, Inc.*, No. 113-2010 CD, 2010 WL 4403285, at *13 (Pa. Ct. Com. Pl. Sept. 9, 2010) (ordering plaintiff to "provide his Facebook and MySpace user names and passwords to counsel" for opposing parties' use).

191. *See In re Grand Jury Subpoena Duces Tecum Dated Mar. 25, 2011*, 670 F.3d 1335, 1341 (11th Cir. 2012).

192. *Privacy in the Digital Age: Encryption and Mandatory Access: Hearing Before the Subcomm. on the Constitution, Federalism, and Prop. Rights of the Comm. on the Judiciary*, 105th Cong. 4 (1998) (statement of Sen. Russell D. Feingold), available at <http://www.loc.gov/law/find/hearings/pdf/00139296461.pdf>.

193. *See Boucher II*, 2009 WL 424718, at *1; *see also Fricosu*, 841 F. Supp. 2d 1232, 1238 (D. Colo. 2012) (ordering defendant to deliver an unencrypted copy of the protected data instead of demanding the password).

194. *See In re Grand Jury Subpoena Duces Tecum Dated Mar. 25, 2011*, 670 F.3d at 1346 (finding that it is not a foregone conclusion that the unencrypted contents contain incriminating evidence when it is not clear that the "[g]overnment knows whether any files

individuals are already concerned about the government accessing communications and communication records through government surveillance,¹⁹⁵ then what will come if the government begins compelling third parties to hand over the trove of information inside one's electronic device? Modern cellphones, with the "vast quantities of personal information" that they contain, could reveal "the privacies of life" for many Americans—not just of the cellphone's owner, but of others as well.¹⁹⁶ After all, many Americans conduct business via their phone, so a search of an individual's phone could include others' confidential information.¹⁹⁷ Thus a broader change is warranted, specifically one that takes into account both privacy and government concerns.

Another potential solution would be for service providers to not collect or store passwords in the first place. Most cellphone manufacturers and telecom providers do not currently collect or store passwords,¹⁹⁸ and any cellphone manufacturers who do collect this information should change direction and not collect or store passwords. An even stronger measure to ensure all service providers take the same action would be legislation prohibiting cellphone manufacturers from collecting and storing passwords. If this approach was adopted, governmental entities would not have the recourse currently available under the SCA to request passwords from service providers. Instead, they would have to obtain this information from individuals who could then invoke the Fifth Amendment. But as long as courts remain split on whether individuals are protected under the Fifth Amendment when the government attempts to compel their passwords, it remains unclear whether individuals' legal rights under the Fifth Amendment exist in this context. And it is unlikely that the government would enact legislation of this type given the government's reaction to Apple and

exist and are located on the hard drives" or that the individual is "even capable of accessing the encrypted portions of the drives").

195. See, e.g., Michael Mimoso, *EFF Makes Case That Fifth Amendment Protects Against Compelled Decryption*, THREATPOST (Oct. 31, 2013, 2:08 PM), <http://threatpost.com/eff-makes-case-that-fifth-amendment-protects-against-compelled-decryption/1027671/>; *NSA Spying on Americans*, EFF, <https://www.eff.org/nsa-spying/> (last visited Oct. 14, 2014).

196. *Riley v. California*, 134 S. Ct. 2473, 2485, 2495 (2014).

197. See *Commonwealth v. Gelfgatt*, 11 N.E.3d 605, 627 (Mass. 1014) (Lenk, J., dissenting) (concluding "that the defendant could not be compelled to enter the decryption key . . . because of the possibility that the computers contain privileged information relating to the defendant's legal clients").

198. "Spokeswomen for Microsoft Corp. and Research In Motion Ltd. say their companies don't collect or store passwords." Angwin, *supra* note 2, at B4; see also Timberg, *supra* note 32 ("Apple once maintained the ability to unlock some content on devices for legally binding police requests but will no longer do so for iOS 8."); Timberg, *supra* note 28 (stating that the next generation of Google's Android operating system will encrypt data by default, thereby preventing anyone but the owner of the device from accessing the phone).

Google's announcements of engineering changes preventing the government from unlocking a user's electronic device.¹⁹⁹

A third—and possibly most promising—solution would be for Congress to amend the ECPA. Congress could change the language of the ECPA and the SCA to allow third parties to invoke these statutes when compelled to hand over passwords, because the Supreme Court has made clear that third parties, even when agents of the suspect, cannot invoke the Fifth Amendment when they receive requests to hand over information that may incriminate the individual target of the investigation. Although Congress must balance the needs of law enforcement with the interests of privacy, this can still be accomplished if a caveat for password production is inserted into the SCA. A proposed amendment would state that third parties are not authorized to disclose an individual's password or encryption key to any other party, nor can they be required to hand over this information to the government unless specifically authorized by the individual account creator, in life-threatening situations, or in exigent circumstances as determined by a court.²⁰⁰ However, given “[o]ur general preference to provide clear guidance to law enforcement through categorical rules,”²⁰¹ Congress would need to create clear standards for the exceptions. Moreover, this caveat would not prevent the government from compelling third parties to disclose the *content* of communications under the ECPA and the SCA. But it would prevent the government from accessing the “great deal of information”²⁰² that lies behind the passwords.²⁰³

B. COMPELLING PASSWORDS FROM EMPLOYERS

As more employers install monitoring software to track employees' every electronic move, the government may soon seek to compel employers to hand over their employees' passwords. Technologies such as keystroke logging make it possible for employers to monitor not only an employee's use of a computer, but also each and every one of their keystrokes.²⁰⁴ In fact,

199. Timberg & Miller, *supra* note 33.

200. *See, e.g., Riley*, 134 S. Ct. at 2494.

201. *Id.* at 2491.

202. *Id.*

203. It is also worth noting that a time may come when computers, cellphones, and other electronic devices use solely biometric information to “unlock” the device instead of a traditional password. Any changes to the laws to protect individuals' privacy rights should account for this broader concept of “passwords.”

204. *See Brahmana v. Lembo*, No. C-09-00106 RMW, 2009 WL 1424438, at *3 (N.D. Cal. May 20, 2009) (“A key logger records each keystroke entered by the user of a particular computer.”); *see also United States v. Ropp*, 347 F. Supp. 2d 831, 837–38 (C.D. Cal. 2004). According to the latest workplace monitoring and surveillance study, 45% of employers track content, keystrokes, and time spent at the keyboard, and 43% store and review

“[k]eystroke logging would provide employers the ability to obtain [an employee’s] passwords if the employee accessed the site using an employer’s computer.”²⁰⁵

Although employers would likely not be considered an ECS²⁰⁶ or RCS²⁰⁷ provider for purposes of the ECPA and the SCA (though one could argue that employers provide their employees with the ability to send and receive electronic communications over their networks), in light of the advancing technology (keystroke logging) and the lagging ECPA, it is critical that Congress address this head on with an amendment to the ECPA. And given that keystroke logging has already been the subject of employment lawsuits,²⁰⁸ employees and employers must understand their privacy rights should the government attempt to compel private employee information from employers under the basis that the employer is an ECS. Specifically, Congress should make clear that an employer (1) is not an ECS or RCS provider and (2) cannot be compelled to hand over employee data via the law of the ECPA and the SCA if the requested information is protected by the Fourth Amendment. Thus, an employer would have to first determine if the employee enjoys a legitimate expectation of privacy in the item being compelled before freely giving it over to law enforcement.

Under a traditional Fourth Amendment analysis, employees should still have a legitimate expectation of privacy in the passwords that they type into an employer-owned device, even if the employer uses keystroke logging software, because this type of software is usually installed in secret. But the

computer files. *The Latest on Workplace Monitoring and Surveillance*, AM. MGMT. ASS’N (Nov. 17, 2014), <http://www.amanet.org/training/articles/The-Latest-on-Workplace-Monitoring-and-Surveillance.aspx>.

205. Alexander Naito, Comment, *A Fourth Amendment Status Update: Applying Constitutional Privacy Protection to Employees’ Social Media Use*, 14 U. PA. J. CONST. L. 849, 864–65 (2012). See also *Brahmana*, 2009 WL 1424438, at *2, *3 (analyzing a situation in which an employer used monitoring tools, such as “Local Area Network Analyzers and keyloggers” to monitor the activities of their employees and in one case intercept an employee’s password to his personal email account). But, keystroke logging “is a risky practice that may violate ECPA,” and is extremely invasive. Corey A. Ciochetti, *The Eavesdropping Employer: A Twenty-First Century Framework for Employee Monitoring*, 48 AM. BUS. L.J. 285, 315, 339 (2011). And employers might want “to revisit how they use keystroke monitoring or logging technology in light of [the N.D. Cal.] ruling.” JACKSON LEWIS, *supra* note 115.

206. 18 U.S.C. § 2510 (2012) (defining electronic communication service as “any service which provides to users thereof the ability to send or receive wire or electronic communications”).

207. *Id.* § 2711 (defining remote computing service as “the provision to the public of computer storage or processing services by means of an electronic communications system”).

208. See, e.g., *Brahmana*, 2009 WL 1424438, at *3 (involving circumstances where the employer recorded the employee “entering his email password ‘using software and hardware monitoring tools such as . . . key loggers”).

employee's reasonable privacy expectation would be compromised if an employer's policy specifically notifies the employee of the use of keystroke logging to track every word typed into an employer-owned device. Aside from this exception, however, passwords should not be treated as voluntarily given to employers in the same way that they are not voluntarily given to service providers. Although employers may have access to this information,²⁰⁹ individuals probably do not assume that their employers need their passwords for any legitimate business purpose, nor do most employees know that employers can make permanent records of the passwords they enter (with of course the exception of a clearly-stated policy holding otherwise).²¹⁰

VII. CONCLUSION

Increased access to rapidly changing technology necessitates Congressional action to ensure individuals' privacy rights remain protected as the Founders intended when they fought their own fight against unrestrained government invasion.²¹¹ The ECPA and the SCA are outdated and do not reflect the privacy concerns that accompany the digital age. Moreover, the Fourth and Fifth Amendments do not adequately protect individuals from forcible and compulsory extortion of an individual's private passwords when the government compels them from a third party—and have even fallen flat in some instances when the government compels passwords from individuals. With all of this in mind, it is incumbent on Congress to consider how the government's recent tactics in compelling users' passwords from third parties flies directly counter to expectations of privacy and denies individuals a meaningful Fifth Amendment privilege when it comes to protecting their password-protected digital footprint.

209. JACKSON LEWIS, *supra* note 115.

210. *See* Reporters Comm. for Freedom of the Press v. Am. Tel. & Tel. Co., 593 F.2d 1030, 1043 (D.C. Cir. 1978) (holding that “[t]o the extent an individual *knowingly* exposes his activities to third parties, he surrenders Fourth Amendment protections”) (emphasis added).

211. *See* Riley v. California, 134 S. Ct. 2473, 2494–95 (2014) (noting that opposition to unrestrained searches “was in fact one of the driving forces behind the Revolution itself . . . for which the Founders fought”).

DISAGREEABLE PRIVACY POLICIES: MISMATCHES BETWEEN MEANING AND USERS' UNDERSTANDING[†]

*Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor,
Brian French, Amanda Grannis, James T. Graves, Fei Liu,
Aleecia McDonald, Thomas B. Norton, Rohan Ramanath,
N. Cameron Russell, Norman Sadeh and Florian Schaub^{††}*

ABSTRACT

Privacy policies are verbose, difficult to understand, take too long to read, and may be the least-read items on most websites even as users express growing concerns about information collection practices. For all their faults, though, privacy policies remain the single most important source of information for users to attempt to learn how companies collect, use, and share data. Likewise, these policies form the basis for the self-regulatory notice and choice framework that is designed and promoted as a replacement for regulation. The underlying value and legitimacy of notice and choice depends, however, on the ability of users to understand privacy policies.

This paper investigates the differences in interpretation among expert, knowledgeable, and typical users and explores whether these groups can understand the practices described in privacy policies at a level sufficient to support rational decision-making. This paper seeks

© 2015 Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, Rohan Ramanath, N. Cameron Russell, Norman Sadeh and Florian Schaub.

[†] For their comments on this study, the authors would like to acknowledge and thank Alessandro Acquisti, Noah A. Smith, and Shomir Wilson, and the participants at the 2014 TPRC 42nd Research Conference on Communication, Information and Internet Policy. Funding for this project was provided, in part, by the National Science Foundation under its Secure and Trustworthy Computing (SaTC) initiative grants 1330596, 1330214, and 1330141 for “TWC SBE: Option: Frontier: Collaborative: Towards Effective Web Privacy Notice and Choice: A Multi-Disciplinary Prospective” and by a Fordham Law School Faculty Research Grant.

^{††} Respectively, Stanley D. and Nikki Waxberg Chair and Professor of Law, Fordham University; Assistant Professor of Computer Science, Carnegie Mellon University; Professor of Computer Science and Engineering & Public Policy, Carnegie Mellon University; Senior Research Programmer, Carnegie Mellon University; Research Fellow, Fordham Center on Law and Information Policy; Ph.D Candidate (Engineering and Public Policy) Carnegie Mellon University; Ph.D Candidate (Computer Science), Carnegie Mellon University; Director of Privacy, Stanford Center for Internet & Society; Privacy Fellow, Fordham Center on Law and Information Policy; Masters Candidate (Computer Science), Carnegie Mellon University; Executive Director, Fordham Center on Law and Information Policy; Professor of Computer Science, Carnegie Mellon University; Postdoctoral Fellow (Computer Science), Carnegie Mellon University.

to fill an important gap in the understanding of privacy policies through primary research on user interpretation and to inform the development of technologies combining natural language processing, machine learning, and crowdsourcing for policy interpretation and summarization.

For this research, we recruited a group of law and public policy graduate students at Fordham University, Carnegie Mellon University, and the University of Pittsburgh (“knowledgeable users”) and presented these law and policy researchers with a set of privacy policies from companies in the e-commerce and news and entertainment industries. We asked them nine basic questions about the policies’ statements regarding data collection, data use, and retention. We then presented the same set of policies to a group of privacy experts and to a group of crowd workers representing typical Internet users.

The findings show areas of common understanding across all groups for certain data collection and deletion practices, but also demonstrate very important discrepancies in the interpretation of privacy policy language, particularly with respect to data sharing. The discordant interpretations arose both within groups and between the experts and the two other groups.

The presence of these significant discrepancies has critical implications. First, the common understandings of some attributes of described data practices mean that semi-automated extraction of meaning from website privacy policies may be able to assist typical users and improve the effectiveness of notice by conveying the true meaning of these policies. However, the disagreements among experts and disagreement between experts and the other groups reflect that ambiguous wording in typical privacy policies undermines the ability of privacy policies to effectively convey notice of data practices to the general public.

The results of this research will, consequently, have significant policy implications for the construction of the notice and choice framework and for the U.S. reliance on this approach. The gap in interpretation indicates that privacy policies may be misleading the general public and that those policies could be considered legally unfair and deceptive. And, where websites are not effectively conveying privacy policies to consumers in a way that a “reasonable person” could, in fact, understand the policies, “notice and choice” fails as a framework. Such a failure has broad international implications since websites extend their reach beyond the United States.

TABLE OF CONTENTS

I.	INTRODUCTION	41
II.	THE LANDSCAPE	42
A.	THE NOTICE AND CHOICE FRAMEWORK	43
B.	RESEARCH ON USABILITY AND TECHNICAL TOOLS	46
1.	<i>Usability</i>	47
2.	<i>Technical Tools</i>	48
a)	P3P	49
b)	Do Not Track.....	50
3.	<i>Research on Automated Understanding of Privacy Policies</i>	51
4.	<i>Unanswered Questions for Automated and Crowdsourced Understanding</i>	52
III.	METHODOLOGY	53
A.	THE PARTICIPANT GROUPS.....	53
B.	PRIVACY POLICY DATA SET	54

C.	PRIVACY POLICY SURVEY AND ANNOTATIONS	56
D.	BACKGROUND DEMOGRAPHICS	61
IV.	DATA COMPARISONS	62
A.	INTRA-GROUP ANNOTATOR AGREEMENT	62
B.	INTER-GROUP ANNOTATOR AGREEMENT	67
C.	QUALITATIVE DATA ANALYSIS	71
1.	<i>Difficulty for Survey Respondents</i>	71
2.	<i>Trends in Selected Text</i>	72
a)	Consensus on Text Selection	72
b)	Interpretation of Policy Silence	74
i)	Data Collection	74
ii)	Data Sharing	75
iii)	Data Deletion	76
iv)	Assumptions for the Interpretation of Specific Textual Language	77
c)	Human Error	80
V.	SIGNIFICANCE OF FINDINGS	83
A.	IMPLICATIONS FOR COMMON UNDERSTANDING AND CONSUMER DECEPTION	83
B.	IMPLICATIONS FOR CROWDSOURCING	85
1.	<i>When Experts Agree</i>	85
2.	<i>When Experts Disagree</i>	86
VI.	CONCLUSIONS	87

I. INTRODUCTION

Privacy policies are verbose, difficult to understand, take too long to read, and may be the least-read items on most websites even as users express growing concerns about information collection practices. But, for all their faults, privacy policies remain the single most important source of information for users to attempt to learn how companies collect, use, and share data. The reason that privacy policies are so important is that the United States takes a “notice and choice” approach to Internet privacy.¹ The idea is that companies post their privacy policies, users read and understand these policies, and then users follow a rational decision-making process to engage only with companies they believe offer an acceptable level of privacy. This structure is designed and promoted as a replacement

1. *See infra* Section II.A.

for regulation. The underlying value and legitimacy of notice and choice thus depends on the ability of users to understand privacy policies.

This paper investigates whether expert, knowledgeable, and typical users can understand the practices described in privacy policies at a level sufficient to support rational decision-making. The paper seeks to fill an important gap in the understanding of privacy policies of typical users through primary research on user interpretation. This research can inform the development of natural language processing and crowdsourcing for policy interpretation and summarization.² Part II of the paper discusses the existing landscape for notice and choice policies and the gaps in prior research on user understanding. Part III then defines the methodology used for the research. Part IV presents the results and reports on discrepancies in the interpretation of privacy policy language among three different groups: privacy experts, law and policy graduate students, and typical users. These results reveal significant discrepancies across the groups. Part V analyzes the critical implications of these discrepancies.

The results of this research will, consequently, have significant policy implications for the construction of the notice and choice framework and for the U.S. reliance on this approach. The implications also expand beyond the United States since websites extend their reach globally.

II. THE LANDSCAPE

This Part will first explain how and why notice and choice is used as a mechanism to address privacy protection. In the United States, notice and choice has become the principal means to address privacy online. While

2. See Norman Sadeh et al., Towards Usable Privacy Policies: Semi-automatically Extracting Data Practices From Websites' Privacy Policies (2014) (poster, ACM Symposium on Usable Security and Privacy (SOUPS)), available at https://cups.cs.cmu.edu/soups/2014/posters/soups2014_posters-paper20.pdf; NORMAN SADEH ET AL., THE USABLE PRIVACY POLICY PROJECT: COMBINING CROWDSOURCING, MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING TO SEMI-AUTOMATICALLY ANSWER THOSE PRIVACY QUESTIONS USERS CARE ABOUT (Carnegie Mellon Univ., Sch. of Computer Sci., Inst. for Software Research Technical Report No. CMU-ISR-13-119, Dec. 2013), available at <http://reports-archive.adm.cs.cmu.edu/anon/isr2013/abstracts/13-119.html>; Steve Bellovin & Sebastian Zimmeck, *Machine Learning Analysis of Privacy Policies*, MICH. TELECOMM. & TECH. L. REV. (forthcoming 2014); Sebastian Zimmeck & Steven M. Bellovin, *Privee: An Architecture for Automatically Analyzing Web Privacy Policies*, in PROCEEDINGS OF THE 23RD USENIX SECURITY SMPOSIUM (Aug. 2014), available at <https://www.usenix.org/conference/usenixsecurity14/technicalsessions/presentation/zimmeck>.

more extensive regulation exists in Europe,³ notice and choice on an international scale plays important roles in the assurance of both privacy rights and international data flows. The implementation of notice and choice in the United States is also affected by the international presence of U.S. websites.

For notice and choice to work effectively, notice must be meaningful for users. This Part will also address prior research into the usability of privacy policies, describe usability problems, and, thus, reveal the gap to be filled by this research.

A. THE NOTICE AND CHOICE FRAMEWORK

Since the 1970s, the United States has promoted fair information practice standards as the guidepost for the protection of privacy.⁴ These principles appear in U.S. law, but the U.S. legal system shies away from comprehensive privacy regulation.⁵ Historically, the United States has addressed discrete privacy issues in narrow statutes, which have been targeted to specific problems and focused on specific actors.⁶ Over the years, the White House, Congress, and the Federal Trade Commission (“FTC”) have encouraged private sector responses to privacy challenges in lieu of new regulation.⁷

Notice and choice are the critical elements for self-regulation of fair information practices. “Notice” is generally described in terms of transparency of the information practices. The FTC has stated the principle as giving

3. See Directive 95/46 of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals With Regard to the Processing of Personal Data and on The Free Movement of Such Data, 1995 O.J. (L 281) (EC).

4. See Robert Gellman, *Fair Information Practices: A Basic History* (Aug. 3, 2014), <http://bobgellman.com/rg-docs/rg-FIPShistory.pdf>.

5. See, e.g., PAUL M. SCHWARTZ & JOEL R. REIDENBERG, *DATA PRIVACY LAW: A STUDY OF UNITED STATES DATA PROTECTION* (1996).

6. See, e.g., Joel R. Reidenberg, *Privacy in the Information Economy: A Fortress or Frontier for Individual Rights?*, 44 FED. COMM. L.J. 195 (1992); SCHWARTZ & REIDENBERG, *supra* note 5.

7. See, e.g., WHITE HOUSE, *CONSUMER DATA PRIVACY IN A NETWORKED WORLD: A FRAMEWORK FOR PROTECTING PRIVACY AND PROMOTING INNOVATION IN THE GLOBAL DIGITAL ECONOMY* (Feb. 23, 2012), *available at* <http://www.whitehouse.gov/sites/default/files/privacy-final.pdf> (voluntary approach); THE PRIVACY PROTECTION STUDY COMMISSION, *PERSONAL PRIVACY IN AN INFORMATION SOCIETY* (July 1977); FEDERAL TRADE COMMISSION, *PRIVACY ONLINE: A REPORT TO CONGRESS 7* (1998), *available at* http://www.ftc.gov/sites/default/files/documents/public_events/exploring-privacy-roundtable-series/priv-23a_0.pdf [hereinafter *PRIVACY ONLINE*]; U.S. DEP’T OF COMMERCE, *PRIVACY AND SELF-REGULATION IN THE INFORMATION AGE* at viii (June 12, 1997), *available at* <http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>.

consumers . . . notice of an entity's information practices before any personal information is collected from them [N]otice of some or all of the following have been recognized as essential to ensuring that consumers are properly informed before divulging personal information:

- identification of the entity collecting the data;
- identification of the uses to which the data will be put;
- identification of any potential recipients of the data;
- the nature of the data collected and the means by which it is collected if not obvious (passively, by means of electronic monitoring, or actively, by asking the consumer to provide the information);
- -whether the provision of the requested data is voluntary or required, and the consequences of a refusal to provide the requested information; and
- -the steps taken by the data collector to ensure the confidentiality, integrity and quality of the data.⁸

Adequate and meaningful notice is necessary for users to be able to make informed decisions about their privacy choices.

“Choice” is typically defined in terms of consent. As the FTC articulates: “[a]t its simplest, choice means giving consumers options as to how any personal information collected from them may be used. Specifically, choice relates to secondary uses of information—*i.e.*, uses beyond those necessary to complete the contemplated transaction.”⁹

Combined, notice and choice are used as a fundamental aspect of privacy protection in the private sector.

Notice and choice is also an important part of the international framework for transborder data flows. In 2000, the European Union and the United States adopted the Safe Harbor agreement to facilitate international data flows.¹⁰ Under the voluntary agreement, U.S. companies would agree to seven principles that were designed to assure the privacy of their E.U. origin

8. See PRIVACY ONLINE, *supra* note 7, at 7–8.

9. *Id.* at 8.

10. See U.S. DEP'T OF COMMERCE, SAFE HARBOR PRIVACY PRINCIPLES (Jul. 21, 2000), available at http://export.gov/safeharbor/eu/eg_main_018475.asp [hereinafter SAFE HARBOR PRINCIPLES]; Joel R. Reidenberg, *E-Commerce and Trans-Atlantic Privacy*, 38 HOUS. L. REV. 717 (2001).

data. The Safe Harbor agreement specifically included “notice” and “choice” as two essential principles.¹¹

The “notice” principle required website operators to “inform individuals about the purposes for which it collects and uses information about them” in “clear and conspicuous language.”¹² The “choice” principle added that those organizations and companies collecting personal information were required to “offer individuals the opportunity to choose (opt out [of]) whether their personal information [was] (a) to be disclosed to a . . . third party or (b) to be used for a purpose that is incompatible with the purpose(s) for which it was originally collected.”¹³ Like the notice principle, “choice” demanded that companies construe their privacy agreements with clarity, stating that “[i]ndividuals must be provided with clear and conspicuous, readily available, and affordable mechanisms to exercise choice.”¹⁴

The global reach of U.S. websites and the impact of foreign standards on American website practices makes the international importance of notice and choice quite high. Similar to the American doctrine, the Article 29 Working Party of European data protection commissioners has looked to notice and choice in a number of initiatives to protect personal data. For example, in 2013, the Working Party released a guidance document for website operators on obtaining website users’ consent for the use of tracking cookies.¹⁵ The Working Party specified that to provide sufficient notice, websites must provide users with specific information about how and why they used cookies.¹⁶ Attaining users’ “blanket consent” without first supplying exact facts would not suffice.¹⁷ The Working Party suggested that website operators configure browsers to require users to actively signify their consent and leave no doubt as to the users’ subjective intent.¹⁸ Moreover, the Working Party emphasized that users be offered a free choice regarding the use of tracking cookies and that users be able to browse a website while declining cookies.¹⁹ The Italian Data Protection Authority internalized the

11. See SAFE HARBOR PRINCIPLES, *supra* note 10.

12. *See id.*

13. *See id.*

14. *See id.*

15. See *Working Document 02/2013 Providing Guidance on Obtaining Consent for Cookies*, The European Commission Article 29 Working Party (adopted Oct. 2, 2013), http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp208_en.pdf.

16. *Id.* at 3.

17. *Id.*

18. *Id.*

19. *Id.* at 5.

Working Party's guidance in May of 2014.²⁰ Among its resolutions was that website banners should contain clear and visible notice and consent requests for users.²¹

Though many of the Working Party's initiatives adopted notice and choice principles, the E.U. data protection authorities also recognize the limitations of notice and choice. At the Safe Harbor Conference in 2009, Dutch Data Protection Authority chairman Jacob Kohnstamm stated in his introductory remarks that enforcement tools (e.g., fines) may be a superior means of ensuring the protection of website users' personal data.²² Kohnstamm stated that "[d]ue to new technological applications[,] transparency alone [notice and choice] is no longer sufficient to guarantee that individuals can oversee the consequences of data processing activities . . . independent oversight is necessary. It is necessary to ensure a level playing field. To ensure that all are abiding to the same rules."²³ In essence, Kohnstamm was expressing important skepticism about the ability for notice and choice to effectively protect individual privacy.

B. RESEARCH ON USABILITY AND TECHNICAL TOOLS

Prior research has shown that the terms contained in policies are frequently unfamiliar to users, and the level of education necessary to understand the policies is high.²⁴ Similarly, research has also shown that notice of privacy policies may not be effective and that some notices are designed to nudge users into disclosing larger quantities of personal information than necessary for the interaction.²⁵ Privacy technologists have

20. See THE ITALIAN DATA PROTECTION AUTHORITY, SIMPLIFIED ARRANGEMENTS TO PROVIDE INFORMATION AND OBTAIN CONSENT REGARDING COOKIES (May 8, 2014), available at <http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/3167654>.

21. *Id.* at 3.

22. Jacob Kohnstamm, Chairman, Dutch Data Protection Authority, Introductory Speech at the Safe Harbor Conference in Washington (2009), available at http://www.dutchdpa.nl/downloads_int/20091118_speech_jko_washington.pdf.

23. *Id.* at 8.

24. See Mark Hochhauser, *Lost in the Fine Print: Readability of Financial Privacy Notices*, PRIVACY RIGHTS CLEARINGHOUSE (July 1, 2001), <https://www.privacyrights.org/ar/GLB-Reading.htm> ("Readability analysis of . . . privacy notices found that they are written at a 3rd-4th year college reading level, instead of the junior high school level that is recommended for materials written for the general public"); Mark A. Graber, Donna M. D'Alessandro & Jill Johnson-West, *Reading Level of Privacy Policies on Internet Health Web Sites*, 51 J. OF FAM. PRAC. 642, 642 (2002), available at http://www.jfponline.com/fileadmin/jfp_archive/pdf/5107/5107JFP_BriefReport.pdf (noting that healthcare websites that contain privacy statements are beyond the reading level of most adults).

25. See Yang Wang et al., *From Facebook Regrets to Facebook Privacy Nudges*, 74 OHIO ST. L.J. 1307 (2013).

developed tools to facilitate notice and choice for online users, but they have achieved only limited success.

1. Usability

Previous work has shown that while users have difficulty finding and using privacy policy information, they remain interested in this information, and when this information is made salient it can impact users' online purchasing decisions.²⁶ Research has also demonstrated that users are interested in several different pieces of information found in privacy policies.²⁷ This suggests that the information in privacy policies could be helpful if presented in a usable way.

Prior research also found that expecting users to read privacy policies places an unreasonably high burden on them because policies take so long to read.²⁸ As a result, there have been several approaches to improving usability. One approach to making privacy policies more accessible is a privacy "nutrition label" that summarizes key points from a privacy policy in a succinct and standard form.²⁹ While this approach has shown promise in research studies, it has not yet been widely adopted.³⁰

26. Janice Y. Tsai, Serge Egelman, Lorrie Cranor & Alessandro Acquisti, *The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study*, 22 INFO. SYS. RESEARCH 254 (JUN. 2011), available at <http://pubsonline.informs.org/doi/pdf/10.1287/isre.1090.0260> (study participants tended to purchase from online retailers who respected their privacy). For the use of default settings to prompt users to both disclose or share personal information, see Issac Dinner, Eric J. Johnson, Daniel G. Goldstein & Kaiya Liu, *Partitioning Default Effects: Why People Choose Not to Choose*, 17 J. EXPERIMENTAL PSYCHOL.: APPLIED 432 (2011); Daniel Goldstein et al., *Nudge Your Customers Toward Better Choices*, 86 HARV. BUS. REV. 12, 99–105 (2008). For a discussion of changes to Facebook's interface that promote sharing, see Fred Stutzman, Ralph Gross & Alessandro Acquisti, *Silent Listeners: The Evolution of Privacy and Disclosure on Facebook*, 4 J. PRIVACY & CONFIDENTIALITY 7 (2012).

27. See Pedro G. Leon et al., *What Matters to Users? Factors that Affect Users' Willingness to Share Information with Online Advertisers*, in PROCEEDINGS OF THE EIGHTH SYMPOSIUM ON USABLE PRIVACY AND SECURITY (SOUPS) (2013), available at http://cups.cs.cmu.edu/soups/2013/proceedings/a7_Leon.pdf; Jialiu Lin et al., *Modeling Users' Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings*, in PROCEEDINGS OF THE ACM SYMPOSIUM ON USABLE SECURITY AND PRIVACY (SOUPS) (July 2014) (quantifying users' willingness to disclose or grant different mobile app privacy permissions—"different pieces of information found in privacy policies").

28. See Alecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 I/S J. L. & POL'Y 540 (2008).

29. Patrick G. Kelley et al., *Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach*, CYLAB (Jan. 12, 2010), http://www.cylab.cmu.edu/research/techreports/2009/tr_cylab09014.html.

30. *Id.*

Layered privacy notices are another approach. Layered notices present a website's privacy policy to users in multiple "layers," with each describing elements of the policy in greater levels of detail and specificity.³¹ Typically, these notices consist of a "short notice in a common template format" coupled with a "longer complete notice."³²

Proponents of this approach advocate that layered notices "easily build consumer trust" and "increase public understanding of privacy and data protection" because the notices are "easy to read and understand."³³ One study revealed, however, that though layered notices enabled study participants to make decisions more quickly, the participants often responded inaccurately to questions about terms they had read in the notices.³⁴ Furthermore, the results suggested that participants rarely probed beyond the initial layer, thus leaving them with "incorrect impressions" of the privacy practices described by the more-complete policy.³⁵

2. *Technical Tools*

Privacy technologists have also developed a variety of tools for users to express privacy preferences and for users to opt out of receiving targeted ads. However, the evaluation and deployment of these technologies over the years shows that the challenges to building effective tools have not been overcome. These challenges include the imposition of burdens on users to understand complex and diverse privacy preferences³⁶ and to comprehend

31. See CTR. FOR INFO. POLICY LEADERSHIP, TEN STEPS TO DEVELOP A MULTILAYERED PRIVACY NOTICE 1 (Mar. 2007), available at http://www.informationpolicycentre.com/files/Uploads/Documents/Centre/Ten_Steps_whitepaper.pdf.

32. *Id.* at 1. This guide proposes that notices contain three layers:

Layer 1 - The short notice: the very minimum, for example, when space is very limited, providing only the identity of the data controller, contact details, and the purposes of processing.

Layer 2 - The condensed notice: covering the basics in less than a page, ideally using subheadings, and covering Scope; Personal information collected; Uses and sharing; Choices (including any access options); Important information; How to contact us.

Layer 3 - The full notice.

Id. at 2–3.

33. *Id.* at 3–4.

34. See generally ALEECIA M. McDONALD ET AL., *A Comparative Study of Online Privacy Policies and Formats*, in PRIVACY ENHANCING TECHNOLOGIES 37 (Aug. 2009), available at <http://robreeder.com/pubs/PETS2009.pdf>.

35. *Id.* at § 6.

36. See, e.g., Michael Benisch et al., *Capturing Location-Privacy Preferences: Quantifying Accuracy and User-Burden Tradeoffs*, 15 J. PERSONAL & UBIQUITOUS COMPUTING 7 (Oct. 2011), available at http://www.normsadeh.com/file_download/142.

general usability features such as mechanisms to opt out in the context of online behavioral advertising.³⁷ These challenges emerge from the most prominent technical efforts to address notice and choice.

a) P3P

Growing concern by Congress and threats from the FTC to regulate online privacy gave rise to the Platform for Privacy Preferences (“P3P”)—a web standard developed by the World Wide Web Consortium (“W3C”) that enables web browsers to read website privacy policies automatically and compare them with user-specified privacy preferences.³⁸ Essentially, P3P would enable users to avoid websites whose practices did not meet their privacy preferences.³⁹ P3P specification 1.0 was launched in 2002.⁴⁰ Though a more developed specification 1.1 working draft was later produced, it was never finalized, as the W3C’s working group “closed . . . due to lack of industry participation” in 2006.⁴¹

While some popular web browsers have integrated P3P tools,⁴² others have not.⁴³ Furthermore, the users of browsers that have integrated P3P are reportedly unaware of the tool.⁴⁴ In addition, thousands of websites that adopted P3P appear to have used P3P codes to circumvent browser cookie blocking, without making accurate computer-readable statements about their

37. Pedro Leon et al., *Why Johnny Can't Opt Out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising*, in PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS 589–98 (2012), available at <http://dl.acm.org/citation.cfm?doid=2207676.2207759>; Pedro Giovanni Leon et al., *What Do Online Behavioral Advertising Privacy Disclosures Communicate to Users?*, in PROCEEDINGS OF THE ACM WORKSHOP ON PRIVACY IN THE ELECTRONIC SOCIETY (WPES) 19–30 (2012), available at <http://doi.acm.org/10.1145/2381966.2381970>.

38. See Lorrie Faith Cranor, *Necessary But Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice*, 10 J. ON TELECOMM. & HIGH TECH. L. 273, 279 (2012).

39. *Id.* See also Kimberly Rose Goldberg, Note, *Platform for Privacy Preferences (“P3P”): Finding Consumer Assent to Electronic Privacy Policies*, 14 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 255 (2003).

40. See Cranor, *supra* note 38, at 279.

41. Cranor, *supra* note 38, at 280.

42. Namely, Microsoft Internet Explorer 6, 7, 8, and 9. See Cranor, *supra* note 38, at 280.

43. Neither Firefox, Safari, nor Chrome have integrated P3P, though “a number of prototype plug-ins and extensions,” “authoring tools,” and “prototype P3P user agents” have been developed. See Cranor, *supra* note 38, at 280–81.

44. Cranor asserts, “While I know of no formal studies, my informal polls of hundreds of audience members at talks I have given suggests that outside of groups of privacy experts, almost nobody has heard of P3P” Cranor, *supra* note 38, at 281.

privacy policies.⁴⁵ Thus, P3P policies have become an unreliable source of privacy policy information.

b) Do Not Track

In 2007, privacy advocates began discussing the creation of a mechanism that would enable users to register their opposition to being tracked online. The mechanism would be similar to the “Do Not Call” list for opting out of telemarketing solicitations.⁴⁶ Over time, this idea developed into a technical mechanism that would allow user agents—including web browsers, cell phones, email clients, and anti-malware packages—to send a “do not track” signal on the user’s behalf.⁴⁷ In 2010, a Federal Trade Commission report requested comments on the idea of Do Not Track (“DNT”).⁴⁸

DNT similarly became a popular topic with legislators. Multiple bills at both the state and federal levels would have made DNT a legal requirement, but only one passed into law: California’s AB 370.⁴⁹ Under AB 370, companies with customers who are California citizens must disclose how, if at all, they respond to an incoming DNT request.⁵⁰ In practice, this disclosure requirement is a de facto national (and international) standard, since most English language websites will likely have at least one visitor from California.⁵¹

By 2012, all major web browsers had implemented an interface for users to send a DNT request. However, the implementation of DNT remains elusive. While DNT is a promising idea, there are three major barriers to

45. PEDRO GIOVANNI ET AL., *Token Attempt: The Misrepresentation of Website Privacy Policies Through the Misuse of P3P Compact Policy Tokens*, in WORKSHOP ON PRIVACY IN THE ELECTRONIC SOCIETY (WPES) 4 (Oct. 2010), available at https://www.cylab.cmu.edu/files/pdfs/tech_reports/CMUCyLab10014.pdf.

46. Christopher Soghoian, *The History of the Do Not Track Header*, SLIGHT PARANOIA (Jan. 21, 2011), <http://paranoia.dubfire.net/2011/01/history-of-do-not-track-header.html>.

47. DONOTTRACK.US, <http://donottrack.us/> (last visited May 19, 2015); J. Mayer et al., *Do Not Track: A Universal Third-Party Web Tracking Opt Out*, INTERNET ENGINEERING TASK FORCE (Mar. 7, 2011), <http://tools.ietf.org/html/draft-mayer-do-not-track-00/>.

48. FED. TRADE COMM’N, A PRELIMINARY FTC STAFF REPORT ON PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: A PROPOSED FRAMEWORK FOR BUSINESSES AND POLICYMAKERS A-4 (December 1, 2010), available at <http://www.ftc.gov/os/2010/12/101201privacyreport.pdf>.

49. Assemb. 370, Ch. 390, Reg. Sess. (Cal. 2013) (amending CAL. BUS. & PROF. CODE § 22575), available at http://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=201320140AB370.

50. *Id.* at § 1(b)(5) (amending CAL. BUS. & PROF. CODE § 22575 (b)(5)).

51. Compliance with AB 370 seems to circumvent the purpose of DNT. An informal survey shows that most of the companies complying with AB 370 offer a vague statement saying that they ignore DNT. These notices are usually contained somewhere in the privacy policy or in a file linked to from the privacy policy.

wide adoption. First, there is no agreement among participants on the treatment of a DNT request by a website, i.e., as to how the website should respond.⁵² Second, only a few prominent companies such as Mozilla, Twitter, and AP News have publicly encouraged DNT, and thus the list of implementers is quite modest.⁵³ Lastly, if there were a new standard that called for all companies to perform a minimum set of actions upon receipt of a DNT signal, companies might simply refuse to allow access to users requesting DNT.

3. *Research on Automated Understanding of Privacy Policies*

Researchers have also considered whether automated processing of privacy policies will be able to provide users with meaningful information for notice and choice.⁵⁴ One recent study explored the possibility of using automated processing and crowdsourcing to interpret website privacy policies.⁵⁵ The study relied on data provided by ToSDR.org, a crowdsourcing project that examined a limited set of privacy policies and that does not use a scientifically-based rating approach for those policies.⁵⁶ The study did, however, find inherent limitations due to “ambiguity of language” and variant human interpretations.⁵⁷

Other studies further investigated the feasibility of leveraging natural language processing and machine learning techniques to tackle the problems of automatic categorization of privacy policies⁵⁸ and grouping segments of policies based on the privacy issues they address.⁵⁹ These studies shed light

52. The World Wide Web Consortium (“W3C”) successfully published a late-stage draft of the technical mechanisms to send and receive DNT signals. See *Tracking Preference Expression (DNT) W3C Last Call Working Draft*, WORLD WIDE WEB CONSORTIUM (Apr. 24, 2014), <http://www.w3.org/TR/tracking-dnt/>.

53. Mozilla published an implementation guide with example source code. See *The Do Not Track Field Guide*, MOZILLA DEVELOPER NETWORK (Oct. 8, 2014), https://developer.mozilla.org/en-US/docs/Web/Security/Do_not_track_field_guide. But, there is no consensus on the treatment of the signal. Several companies have announced they honor DNT. See *Do Not Track: Implementations*, DONOTTRACK.US, <http://donottrack.us/implementations/> (last visited May 19, 2015).

54. See Sadeh et al., *supra* note 2; Bellovin & Zimmeck, *supra* note 2; Zimmeck & Privee, *supra* note 2.

55. Bellovin & Zimmeck, *supra* note 2; Privee, *supra* note 2.

56. See *supra* note 55.

57. *Id.*

58. Waleed Ammar et al., *Automatic Categorization of Privacy Policies: A Pilot Study*, CARNEGIE MELLON UNIV., SCH. OF COMPUTER SCIENCE, TECHNICAL REPORT NO. CMU-ISR-12-114, CMU-LTI-12-019, (2012).

59. Fei Liu et al., *A Step Towards Usable Privacy Policy: Automatic Alignment of Privacy Statements*, in PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS (COLING) (Aug. 2014); Rohan Ramanath et. al, *Unsupervised Alignment of Privacy*

on automatic methods of understanding privacy policies; however, it is not clear if the existing natural language techniques are able to fully decode the sophistication and ambiguity of privacy policies. A more promising approach will likely involve combining such techniques with machine learning and crowdsourcing, hence the importance of this study.

Another study examined the manual translation of privacy policies into a specialized mathematical logic.⁶⁰ The study results include heuristics for mapping variant interpretations into a single, canonical representation expressed in logic and a demonstration of how this logical representation can be used to answer questions about information collection, use, and sharing. For example, one heuristic includes mapping certain verbs, such as “transfer,” “share,” and “access” to events in which a data holder shares personal information with a third party.⁶¹ In particular, the verb “access” is ambiguous because it can map to collection, use, or sharing depending on the stakeholder viewpoint, i.e., who has access. Other ambiguities, such as omissions, generic terms, and terms that have varying technical interpretations can lead to variant interpretations, some of which may be unintended by the policy authors. While the formalization does enable automated reasoning to detect policy conflicts due to ambiguous policy statements, it requires special training to perform the translation into logic. As with any policy document, the logical representation must also be maintained as the natural language policy changes. This prior work leaves open the question of how automated processing and crowdsourcing might function on an enormously broad set of privacy policies with a systematic approach to rating those policies. Although the preliminary results⁶² are promising, it is not clear from this prior work whether a level of automation can be reached that would enable the process to be conducted on a web scale.

4. *Unanswered Questions for Automated and Crowdsourced Understanding*

In light of the present state of research, this study tests the comprehension and clarity of privacy notices on a larger scale, with the aim

Policies using Hidden Markov Models, in PROCEEDINGS OF THE ANN. MEETING OF THE ASSOC. FOR COMPUTATIONAL LINGUISTICS (ACL) (June 2014).

60. Travis Breaux et. al., *Eddy*, *A Formal Language for Specifying and Analyzing Data Flow Specifications for Conflicting Privacy Requirements*, 19 REQUIREMENTS ENG'G J. 281, 281–307 (Dec. 2013).

61. *Id.*

62. Travis Breaux & Florian Schaub, *Scaling Requirements Extraction to the Crowd: Experiments on Privacy Policies*, in 22ND IEEE INT'L REQUIREMENTS ENG'G CONFERENCE (2014), available at <http://www.cs.cmu.edu/~breaux/publications/tdbreaux-re14.pdf>.

of making a prognosis about the viability of large-scale semi-automated analysis and review of privacy policies. Prior work shows that policy ambiguity may challenge the ability of natural language processing to be effective. Crowdsourcing may not fully remedy these ambiguities, but crowdsourcing might help overcome some of these limitations depending on how far the interpretation of non-ambiguous elements might be scaled (e.g., does it always require expert annotators and could one leverage the so-called “wisdom of the crowds”?). Accordingly, this study is designed to explore the clarity of privacy policies in more detail by examining how three groups with different levels of expertise understand privacy notices. The goal is to elicit commonalities and differences in the comprehension and interpretation of websites’ privacy policies across groups of participants with varying legal backgrounds and training.

III. METHODOLOGY

The research methodology was designed to discover how three different user groups would each interpret specific language in privacy policies. As discussed below, the participant groups were chosen to reflect expert, knowledgeable, and crowd workers representing typical users. Privacy policies were systematically collected from the web, and a survey was created to probe user understanding of the policies. In addition, background information was collected from the survey respondents.

A. THE PARTICIPANT GROUPS

Three groups participated in this study: 1) crowd workers representing typical users; 2) knowledgeable users; and 3) privacy policy experts. These groups were recruited as follows:

- 1) *Crowd workers* were recruited on Amazon Mechanical Turk (“MTurk”) as a representative sample of the general population.⁶³ MTurk is an internet marketplace that uses human intelligence by crowdsourcing the performance of tasks requested by individuals or business.⁶⁴ The data used for this study came from twenty-eight crowd workers.⁶⁵ Previous studies have shown that MTurk provides a suitable participant pool for conducting research studies

63. MTurk is an Amazon website that pays users to complete proposed tasks.

64. See AMAZON MECHANICAL TURK (BETA), <http://aws.amazon.com/mturk/> (last visited May 19, 2015).

65. As discussed in *infra* Sections III.B & C, five crowd workers annotated each of the six policies. Of the 28 crowd workers, 27 annotated only one of the six policies and 1 crowd worker annotated 3 of the six policies. The 1 crowd worker who annotated multiple policies annotated the Barnes & Noble, ABC News, and Washington Post policies.

and that the demographic distribution of crowd workers on MTurk is comparable to the general U.S. population.⁶⁶ These workers were paid \$6.00 per reviewed policy. To be eligible to participate, these workers were required to have at least a 95% approval rating for 500 completed tasks on MTurk and be U.S. residents.⁶⁷ U.S. residency was verified with a question asking about the worker's country of residence. Multiple screening checks were applied in order to determine whether a crowd worker made an honest effort in completing the task. This vetting consisted of checking for the duration spent on the task, whether question responses were accompanied by meaningful text selections (see below), and whether the participant provided actual words for the answers to a Cloze test, a test that requires participants to replace several missing words in a piece of text to assess their general reading comprehension. All crowd worker submissions satisfied these checks, likely because the required qualification (95% approval rating on 500 tasks) and the relatively high pay (\$6.00) were sufficient to motivate honest participation in the study.

2) *Knowledgeable users* consisted of five graduate students with a background in law, public policy, or computer science who were recruited from Fordham University, Carnegie Mellon University, and the University of Pittsburgh. These five knowledgeable users were hired as research assistants.

3) *Privacy policy experts* consisted of four of the study authors who are experienced law and public policy scholars. The purpose of these expert annotations was to determine the degree of agreement between experts, as well as to investigate the deviation of professional interpretation from the interpretation by knowledgeable users and crowd workers.

B. PRIVACY POLICY DATA SET

We collected unique privacy policies from the top websites ranked by Alexa.com, the most prominent measurement company for web traffic data.⁶⁸

These policies were collected during a period of six weeks between

66. See Tara S. Behrend et al., *The Viability of Crowdsourcing for Survey Research*, 43 BEHAVIOR RES. METHODS 800, 800–13 (2011); Gabriele Paolacci et al., *Running Experiments on Amazon Mechanical Turk*, 5 JUDGMENT & DECISION MAKING 411, 411–19 (2010); Gabriele Paolacci & Jesse Chandler, *Inside the Turk: Understanding Mechanical Turk as a Participant Pool*, 23 CURRENT DIRECTIONS IN PSYCHOL. SCI. 184, 184–88 (2014).

67. The MTurk rating level was set to assure that workers would take the task seriously and the U.S. residency requirement was set to assure that workers would not assume rights that exist in foreign countries.

68. Alexa ranks the popularity of websites based on their traffic. See *About Us*, ALEXA (2014), <http://www.alexa.com/about/>.

December 2013 and January 2014. They provide a snapshot of privacy policies from mainstream websites based on Alexa.com's website categories.⁶⁹ Since locating a website's policy is not a trivial task, we crowdsourced the privacy policy document collection using MTurk.⁷⁰

In an earlier exploratory study, fifteen websites were selected from each of the Alexa.com "news" and "shopping" categories and were used for initial crowdsourcing analysis. The websites were selected in a top-down fashion using the rankings provided by Alexa.com. Additionally, two websites (amazon.com, yahoo.com) were set aside as a development data set and used for testing the crowdsourcing interface.

In this study, we focus on U.S commercial websites. From the policy data set, three privacy policies were manually selected from the "news" category and three policies were manually selected from the "shopping" category. The main websites for the companies whose privacy policies were selected for study are listed below along with the date of each policy's last revision as of the moment of collection:

News sites:

- ABC News (December 30, 2013):
<http://abcnews.go.com/>
- Washington Post (November 15, 2011):
<http://www.washingtonpost.com/>

69. Of the seventeen categories, two were excluded: the "Adult" and the "World" category. The "World" category was excluded since it contained mainly popular websites in different languages, and we opted to focus on policies in English in this study.

70. Though many well-regulated commercial websites provide a "privacy" link on their homepages, not all do. Neither is there a standardized URL format for privacy policies. Even once the policy's URL is identified, extracting the policy text presents the usual challenges associated with scraping documents from the web. Since every site is different in its placement of the document (e.g., buried deep within the website, distributed across several pages, or mingled together with Terms of Service) and format (e.g., HTML, PDF, etc.), and since we aimed to preserve as much document structure as possible (e.g., section labels), full automation was not a viable solution. For each website, we created a "human intelligence task" or HIT in which a worker was asked to copy and paste the following privacy policy-related information into text boxes: (i) privacy policy URL; (ii) last updated date (or effective date) of the current privacy policy; (iii) privacy policy full text; and (iv) the section subtitles in the top-most layer of the privacy policy. To identify the privacy policy URL, workers were encouraged to go to the website and search for the privacy link. Alternatively, they could form a search query using the website name and "privacy policy" (e.g., "Amazon.com privacy policy") and search in the returned results for the most appropriate privacy policy URL. Each HIT was completed by three workers who were each paid \$0.05 per HIT. The collected privacy policies were further validated through manual review by one of the authors to ensure quality annotations.

- Weather Underground (October 30, 2013):
<http://www.wunderground.com/>

Shopping sites:

- Barnes and Noble (May 7, 2013):
<http://www.barnesandnoble.com/>
- Lowe's (April 25, 2013):
<http://www.lowes.com/>
- Overstock (January 9, 2013):
<http://www.overstock.com/>

These six privacy policies are included in Appendix A.

C. PRIVACY POLICY SURVEY AND ANNOTATIONS

The study focused on three key privacy policy elements: the collection of information, sharing of information, and deletion of information. These were chosen to reflect important user concerns and were selected based on an analysis of FTC privacy enforcement actions, which identified surreptitious collection, unauthorized disclosure, and wrongful retention of personal information as the most significantly contested online information practices.⁷¹ The study asked about four information types shown to be highly relevant to users in previous studies.⁷² These information types were: contact information (such as an address), financial information (such as payment information), current location information, and health information.

To discover commonalities and differences in interpretation between our different participant groups, we created a survey for participants that asked nine questions about different data practices described in a website's privacy

71. See JOEL R. REIDENBERG, ET AL., FORDHAM CTR. ON LAW & INFO. POL'Y, PRIVACY ENFORCEMENT ACTIONS (June 24, 2014), http://law.fordham.edu/assets/CLIP/CLIP_Privacy_Case_Report_-_FINAL.pdf [hereinafter PRIVACY ENFORCEMENT ACTIONS]. A fourth aspect—inadequate security for personal information—was not considered in this study, because privacy policies often contain only vague statements on security measurements.

72. See Mark S. Ackerman, Lorrie Faith Cranor & Joseph Reagle, *Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences*, in PROCEEDINGS OF THE 1ST ACM CONFERENCE ON ELECTRONIC COMMERCE (EC), 1–8 (1999); Adam N. Joinson, Ulf-Dietrich Reips, Tom Buchanan & Carina B. P. Schofield, *Privacy, Trust, and Self-Disclosure Online*, 25 HUMAN-COMPUTER INTERACTION 1 (2010); Craig E. Wills & Mihajlo Zeljkovic, *A Personalized Approach to Web Privacy: Awareness, Attitudes and Actions*, 19 INFO. MGMT. & COMP. SECURITY 53 (2011); Pedro G. Leon, et al., *What Matters to Users?: Factors That Affect Users' Willingness to Share Information With Online Advertisers*, in PROCEEDINGS OF ACM SYMPOSIUM ON USABLE SECURITY AND PRIVACY (SOUPS) 7 (2013).

policy. These questions (four collection questions, four sharing questions, one deletion question) are described below.

Each study participant was asked to answer the set of survey questions for each of the respective policies. For each answer, the participant was asked to select the text from the policy sections corresponding to the chosen answer. Each of the experts annotated the same set of six privacy policies specified above. While each of the knowledgeable users and crowd workers annotated more than six policies, these six policies reviewed by the experts were the only policies considered among those surveyed for the other participants.⁷³ The annotation process was completed using an online tool created for the task. Participants would select sentences and text passages in the policy with the mouse and then add those passages into a text field under the question by clicking a button. Participants could add one or multiple policy statements for their answers. All answer responses other than the *not applicable* response option required the selection of at least one accompanying text segment.⁷⁴

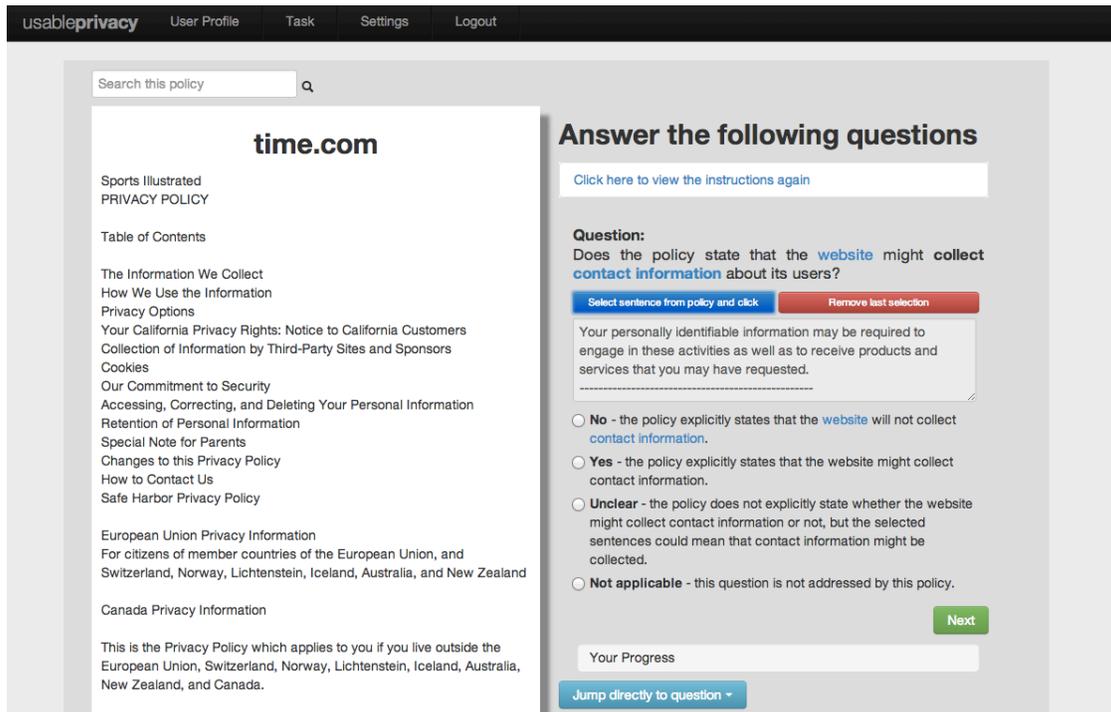
The annotation tool and the wording of questions and response options were refined over multiple iterations of pilot testing and an exploratory experiment. The pilot experiment used six participants (law and computer science graduate students) who each annotated fifteen policies and provided feedback in semi-structured interviews.

73. MTurk crowd workers could choose to annotate only one policy or multiple policies and were compensated separately for each. Crowdsourcing tasks were created so that we obtained at least five annotations from different crowd workers per policy. The majority of MTurk crowd workers chose to annotate only a single policy. Each knowledgeable user annotated twenty- privacy policies in total. We further conducted semi-structured interviews with all five knowledgeable users to gain deeper insights into their annotation strategies and their interpretation of our elicitation questions and policy statements.

74. The online annotation tool further provided participants with detailed instructions on how to complete the annotation task. Participants were instructed to answer questions only for the company's main website and ignore privacy policy statements pertaining to other aspects of a company's business, such as mobile applications, physical stores, or other websites operated by the same company. Participants were further asked to ignore statements pertaining to a specific subset of users, such as statements addressing California privacy laws, E.U. Safe Harbor regulation, or the Children's Online Privacy Protection Act. The instructions further clarified that the most fitting option should be selected based on the information given in the shown privacy policy and that "unclear" should be selected if multiple options would seem to apply, statements are ambiguous or contradictory, or if access to additional linked policies (e.g., a separate cookie policy) would likely be required to answer a question conclusively. We also provided definitions for common terms in the questions and response options (see blue highlights in Figure 1) for further clarification.

The final version of the online tool is illustrated in Figure 1 below. The scrollable privacy policy is displayed on the left side of the screen, and one question is shown at a time in a sidebar on the right. Participants could either progress through the questions sequentially or jump between questions in order to enable participants to quickly translate discovered policy statements into responses to our survey questions.

Figure 1
Online tool for privacy policy annotations.



display a tooltip with a definition of the respective term.

The survey questions on collection of personal information (Q1–Q4) inquired whether contact information (Q1), financial information (Q2), current location information (Q3), or health information (Q4) is being collected by the given website. Participants could choose between four answer options:

Answer Option 1: No—the policy explicitly states that the website will not collect [specified type of information (i.e., contact, financial, etc.)].

Answer Option 2: Yes—the policy explicitly states that the website might collect [specified type of information (i.e., contact, financial, etc.)].

Answer Option 3: Unclear—the policy does not explicitly state whether the website might collect [specified type of information (i.e., contact, financial, etc.)] or not, but the selected sentences could mean that [specified type of information (i.e., contact, financial, etc.)] might be collected.

Answer Option 4: Not applicable—this question is not addressed by this policy.

While the *Yes* and *No* options capture explicit statements in the policy, the *Unclear* option enabled participants to note ambiguity in the policy regarding the collection of a specific information type. The *Not applicable* option, on the other hand, allowed for distinguishing between a policy containing ambiguous statements or no statement at all.

The questions on sharing of personal information (Q5–Q8) inquired whether a website would share contact information (Q5), financial information (Q6), current location information (Q7), or health information (Q8) with third parties. If the policy stated that personal information would be shared with third parties, participants could indicate whether the information would be shared for the purpose of fulfilling a core service (e.g., payment processing or delivery of purchased goods), for purposes other than core services, or for purposes other than core services but only with explicit consent. Hence, the sharing of personal information questions offered six response options:

Answer Option 1: No sharing—the policy explicitly states that the website will not share [specified type of information (i.e., contact, financial, etc.)] with third parties.

Answer Option 2: Sharing for core service only—the policy explicitly states that the website might share [specified type of information (i.e., contact, financial, etc.)] with third parties, but only for the purpose of providing a core service, either with explicit or implied consent/permission from the user.

Answer Option 3: Sharing for other purpose—the policy explicitly states that the website might share [specified type of information (i.e., contact, financial, etc.)] with third parties for other purposes. The policy makes no statement about the user's consent/permission or user consent is implied.

Answer Option 4: Sharing for other purpose (explicit consent)—the policy explicitly states that the website might share [specified type of information (i.e., contact, financial, etc.)] with third parties for a purpose that is not a core service, but only if the user provided explicit permission/consent to do so.

Answer Option 5: Unclear—the policy does not explicitly state whether the website might share [specified type of information (i.e., contact, financial, etc.)] with third parties or not, but the selected sentences could mean that *contact information* might be shared with third parties.

Answer Option 6: Not applicable—this question is not addressed by this policy.

The question on deletion of personal information (Q9) asked about the website's respective deletion policy statements. We explicitly excluded any statements concerning retention for legal purposes, as we sought to assess policy statements that relate to issues of wrongful retention of personal information.⁷⁵ If the policy explicitly stated that information could be removed, participants could indicate whether information would be removed fully or whether some or all of the information may be retained for other purposes:

Answer Option 1: No removal—the policy explicitly states that the user will not be allowed to delete their personal data.

Answer Option 2: Full removal—the policy explicitly states that users may delete their personal data and that no data will be retained for any purpose, whether the data was provided directly by the user, generated by the user's activities on the website, or acquired from third parties.

Answer Option 3: Partial removal—the policy explicitly states that users may delete their personal data but some/all of the data might be retained for other purposes, whether the data was provided directly by the user, generated by the user's activities on the website or acquired from third parties.

Answer Option 4: Unclear—the policy does not explicitly state whether users may delete their personal data or not (e.g., it only talks about editing information).

Answer Option 5: Not addressed—this question is not addressed by this policy.

After answering all nine annotation questions for a given privacy policy, participants were shown an additional screen with three questions asking them whether they had ignored parts of the privacy policy because they did not refer to the company's main website (yes/no), whether the privacy policy contained pointers or links to other policy documents

75. See PRIVACY ENFORCEMENT ACTIONS, *supra* note 71.

(yes/no), and to rate “how easy or difficult it was to answer the previous nine questions for this privacy policy” on a five-point Likert scale (from “very difficult” to “very easy”).

D. BACKGROUND DEMOGRAPHICS

Additionally, after completing the privacy policy annotations, participants were further asked to complete a background questionnaire, which consisted of multiple parts. First, participants were asked to rate their ability to understand legal texts on a five-point scale (from “very difficult” to “very easy”). This self-assessment was followed by questions about the level of legal training they had received and whether they worked in a position that required legal expertise. The second part collected basic demographic information, namely, gender, education level, and primary occupation. In the third part, participants were presented with a Cloze test—a test that requires participants to replace several missing words in a piece of text—to assess their general reading comprehension.⁷⁶ The background questionnaire closed with a number of questions about the annotation experience. Participants were asked to rate the perceived ease or difficulty of answering each of the nine annotation questions on a seven-point scale (from “very difficult” to “very easy”) and the helpfulness of provided instructions and definitions on a seven-point scale (from “not at all helpful” to “very helpful”). Lastly, an open-ended question allowed participants to further comment on difficulties with terms, questions, or answer options.

In terms of gender and age, the twenty-eight crowd workers who annotated the six relevant privacy policies were 64% male (18) and 36% female (10). They ranged in age from 22 to 63 (with a median age of 29). Our group of five knowledgeable users was 40% male (2) and 60% female (3). Their ages were slightly younger (23 to 35 years of age with a median age of 24). The four privacy policy experts were 75% male (3) and 25% female (1). They were slightly older in comparison (34 to 53 years with a median age of 42).

With respect to education, all four experts have a graduate degree, and all five knowledgeable users have at least a bachelors degree (one has a graduate degree). The crowd workers were less educated, with only 46% having a bachelors degree or higher 2.

76. See CAMBRIDGE ENGLISH PROFICIENCY CERTIFICATE OF PROFICIENCY IN ENGLISH CEFR LEVEL C2, HANDBOOK FOR TEACHERS, UNIV. OF CAMBRIDGE 14 (2013), available at https://www.teachers.cambridgeesol.org/ts/digitalAssets/117848_Cambridge_English_Proficiency__CPE__Handbook_2013.pdf.

Crowd workers' primary occupations were diverse and included administrative support (6), service industry (4), unemployed (4), art/writing/journalism (3), business/management/financial (3) student (2), art/writing/journalism (3), business/management/financial (3), education (2), and other (4). None of the crowd workers selected "legal (e.g., lawyer, law clerk)" as a primary occupation. Eighteen crowd workers indicated they had no legal training at all. Seven indicated that they had no legal training but that their background in another field provided them with some legal experience. Finally, three indicated that they were knowledgeable in legal matters but had no formal training. Privacy policy experts were all researchers and scholars, and two of them studied law. Knowledgeable users were all current students of law, computer science, or public policy.

IV. DATA COMPARISONS

This part reports on the data collected as a result of the methodology described above. The part will first analyze the empirical data on intra-group annotator agreement. Then it will analyze the empirical data on inter-group annotator agreement. As will be shown, high levels of agreement within a group ("intra-group" agreement) does not guarantee that this group converged on the same answers as other groups ("inter-group" agreement). The part will conclude with a discussion of qualitative trends observed in the data.

A. INTRA-GROUP ANNOTATOR AGREEMENT

The degree of agreement within each user group is shown in Tables 1–3.⁷⁷ Table 1 shows intra-group agreement for all questions and over the six policies annotated by the experts group for data collection, while Tables 2 and 3 show intra-group agreement for data sharing and deletion respectively. Intra-group agreement is measured by the median level of group member agreement on the same answer across all the policies.⁷⁸ First, the most frequently chosen answer (mode) for each question in each policy was identified. Then, to determine the level of agreement with the mode answer, the percentage of annotators selecting that mode answer was calculated for

77. The tables do not distinguish between the website policies of news and shopping sites. The agreement rate among annotators was almost identical between the news and shopping categories for knowledgeable users. Crowd workers had similar agreement rates to the knowledgeable users on news sites, but slightly less agreement on the policies for shopping sites.

78. Because of the small number of annotators, mean and standard deviation calculations would not provide an accurate representation of group member consensus.

each question in each policy. The median level of agreement across all policies for the same question was calculated to reflect the group consensus on an answer choice. A median value of 100% means that all group members agree on the same answer choice for the survey question for at least four out of the six policies and share the same understanding of those privacy policies. As the median agreement declines, the group members understand some of the policies differently from each other. Any value less than 100% reflects that the annotator group had no consensus answer on the same question for three or more policies—the lower the value, the greater the disagreement. The tables also reflect the median level of agreement when related answer choices were combined.⁷⁹ By combining answer choices, we can determine if there is at least a consensus on the way the policy broadly treats information. When answer choices are combined, the table shows the median level of agreement for each answer combination. Differences in the mode answer choice across the policies are reflected by median calculation based on those policies with the same mode.⁸⁰ The difference between the median level of agreement across all answer choices and the median agreement when several answer choices are combined reflects that group members recognize that the policy addresses a particular point, but they do not share the same understanding of the nuances.

Complete agreement by all members within each of the groups was uncommon and, as shown in the tables, agreement was not distributed evenly across questions. For data collection as reflected in Table 1 below, experts had a consensus on contact and location information (100% median level of agreement across all policies), but they varied in their understanding of the collection of financial and health information (87.5% and 50% respectively). Knowledgeable users shared the same understanding on the collection of contact and financial information (100% median level of agreement across all policies), but not location (70%) or health information (80%). Interestingly, the knowledgeable users had greater agreement on health information than

79. In Table 1, we combined the answer options “Unclear” and “Not applicable” as they were not differentiated consistently by all annotators. In Table 2, we combined answer choices 2–4 (“Sharing for core service only,” “Sharing for other purpose,” and “Sharing for other purpose (explicit consent)”), as all three of them describe that sharing with third parties is taking place, but differentiate between consent models; we further combined answer choice 5–6 (“Unclear” and “Not applicable”) for the same reasons as above. In Table 3, we combined answer choices 2–3 (“full removal” and “partial removal”) as they describe that removal is possible but vary in whether data is retained; as well as answer choices 4–5 (“Unclear” and “Not applicable”) for the same reasons as above. *See infra* Tables 1–3.

80. This means the median score for a given answer choice combination may be based on fewer than six policies and is independent from the median calculations of the other answer choice combinations.

the experts. These results likely indicate that knowledgeable users missed important ambiguity in the privacy policy. Lastly, crowd workers had the lowest level of shared understanding compared to the other groups for contact information (90%), location information (90%), financial information (50%), and health information (70%). The crowd workers had a higher level of agreement on location information than the knowledgeable users, though less than the experts.

Table 1
Data Collection: Intra-group Agreement

Level of Agreement on the Same Answer (Median Across All Policies)	Collect Contact	Collect Financial	Collect Location	Collect Health
Experts				
All choices	100 %	87.5 %	100 %	50 %
Answer Choice 1	n/a	n/a	n/a	n/a
Answer Choice 2	100 %	100 %	100 %	n/a
Answer Choice 3-4	n/a	75 %	n/a	100 %
Knowledgeable Users				
Using All Answers	100 %	100 %	70 %	80 %
Answer Choice 1	n/a	n/a	n/a	n/a
Answer Choice 2	100 %	100 %	90 %	n/a
Answer Choice 3-4	n/a	90 %	100 %	100 %
Crowd Workers				
Using All Answers	90 %	50 %	90 %	70 %
Answer Choice 1	n/a	n/a	n/a	n/a
Answer Choice 2	90 %	100 %	90 %	n/a
Answer Choice 3-4	n/a	60 %	n/a	100 %

For data sharing, Table 2 (shown below) indicates that the experts did not have a high level of consensus on the meaning of the privacy policies (ranging from 75% median level of agreement on contact information across all policies to 50% on health data). However, the level of agreement

improves when the disclosure choices are aggregated (i.e., when answer choices 2–4 are collapsed into one). This means that the experts recognize, though never unanimously, that sharing occurs, but they disagree as to the conditions for sharing as set out in the privacy policies. The knowledgeable users had weak agreement on their interpretations of sharing for contact information (60% median level of agreement across all policies), financial information (50%), and location information (60%). Oddly, the knowledgeable users' level of agreement for health information across all policies was greater than that for the experts (80% median level vs. 50% median level). This means that the knowledgeable users did not perceive as much ambiguity as the experts and suggests that the knowledgeable users may have misunderstood the privacy policies' sharing terms for health information. The crowd workers similarly had weak agreement on their interpretation of policy statements on sharing and, similarly, had a greater consensus on the sharing of health information as compared to the experts.

Table 2

Data Sharing: Intra-group Agreement

Level of Agreement on the Same Answer (Median Across All Policies)	Share Contact	Share Financial	Share Location	Share Health
Experts				
All Choices	75 %	62.5 %	62.5 %	50 %
Answer Choice 1	n/a	n/a	n/a	n/a
Answer Choice 2–4	87.5 %	87.5 %	75 %	n/a
Answer Choice 5–6	n/a	n/a	75 %	75 %
Knowledgeable Users				
All Choices	60 %	50 %	60 %	80 %
Answer Choice 1				
Answer Choice 2–4	80 %	80 %	100 %	n/a
Answer Choice 5–6	60 %	60 %	100 %	100 %
Crowd Workers				
All Choices	60 %	60 %	40 %	80 %
Answer Choice 1	n/a	n/a	n/a	n/a
Answer Choice 2–4	100 %	60 %	60 %	n/a
Answer Choice 5–6	n/a	100 %	70 %	100 %

For deletion, Table 3 below shows that the experts were not in complete agreement on the specific terms for data deletion (75% median level of agreement across all policies). They did, however, agree (100%) on whether some deletion options were available when the full and partial deletion options were aggregated. Knowledgeable users had less agreement than the experts on the terms of deletion policies (60%), but like the experts, when full and partial deletion choices were combined, the knowledgeable users had a complete consensus that at least some deletion was possible (100%). Crowd workers had the least agreement on data deletion (50%), but also reached complete consensus on policies that allowed at least partial deletion (100%).

Table 3

Data Deletion: Intra-group Agreement

	Level of Agreement on the Same Answer (Median Across All Policies)
Experts	
All Choices	75 %
Answer Choice 1	n/a
Answer Choice 2-3	100 %
Answer Choice 4-5	n/a
Knowledgeable Users	
Using All Answers	60 %
Answer Choice 1	n/a
Answer Choice 2-3	100 %
Answer Choice 4-5	n/a
Crowd Workers	
Using All Answers	50 %
Answer Choice 1	n/a
Answer Choice 2-3	100 %
Answer Choice 4-5	n/a

The lack of complete agreement within each group was further confirmed by the Fleiss' Kappa (K) statistic. The Fleiss' Kappa statistic is used to compute the proportion of agreement above chance for multiple

raters who independently rate the same number of items.⁸¹ Inter-rater reliability statistics that account for chance agreement factor in the probability distribution that answer choices are selected across the rating sample. If there is a high probability that an answer choice will be chosen across items, then there is a higher probability that raters will select this option by chance alone. Landis and Koch proposed arbitrary divisions for Kappa that rank 0.41–0.60 as moderate agreement and 0.61–0.80 as substantial agreement.⁸² The computed Fleiss' Kappa for each group of raters are as follows: Expert $K = 0.496$, Knowledgeable Users $K = 0.703$, and Crowd Workers $K = 0.500$.

B. INTER-GROUP ANNOTATOR AGREEMENT

The shared understandings across user groups are shown in Tables 4–6. Table 4 shows the inter-group agreement levels for data collection. Table 5 reflects the level of inter-group agreement on data sharing. Table 6 presents the interpretation of data deletion. The tables were calculated using combined answer choices to determine whether there were shared understandings of the broad practices.⁸³ Inter-group agreement was measured across all the policies by comparing the answer choices of the knowledgeable users and crowd workers with the mode answer choices of the experts.⁸⁴

81. See Joseph Fleiss, *Measuring Nominal Scale Agreement Among Many Raters*, 76 PSYCHOL. BULL. 378 (1971).

82. See J. Richard Landis & Gary G. Koch, *The Measurement of Observer Agreement for Categorical Data*, 33 BIOMETRICS 159 (1977).

83. In Table 4, we combined the answer options “Unclear” and “Not applicable” as they were not differentiated consistently by all annotators. In Table 5, we combined answer choices 2–4 (“Sharing for core service only,” “Sharing for other purpose,” and “Sharing for other purpose (explicit consent)”) as all three of them describe that sharing with third parties is taking place, but differentiate between consent models; we further combined answer choice 5–6 (“Unclear” and “Not applicable”) for the same reasons as above. In Table 6, we combined answer choices 2–3 (“full removal” and “partial removal”) as they describe that removal is possible but vary in whether data is retained; as well as answer choices 4–5 (“Unclear” and “Not applicable”) for the same reasons as above. See *infra* Tables 4, 6.

84. One caveat to the analysis is that the consensus category is defined as the median level of agreement. For statisticians, determining the level of agreement based on consensus category has two potential drawbacks. First, consensus can be based on the mode, median, or mean agreement level, which makes it difficult to compare results across studies that use a different basis for consensus. Second, there are different response patterns that can lead to the same consensus level and same conclusion, despite different variability in the responses (i.e., the consensus technique is not sensitive to variability). Another way to examine consensus comes from Vanbelle and Albert who describe a novel agreement index that is a natural extension of Cohen's Kappa for measuring the overall proportion of agreement above chance between two groups of raters that also accounts for the heterogeneity of each group. See S. Vanbelle & A. Albert, *Agreement Between Two Groups of Raters*, 74 PSYCHOMETRIKA 477 (2009). The Vanbelle/Albert Kappa for collection, sharing and

First, to establish the control values, the experts' most frequently chosen answer on each question for each policy was identified. The level of agreement among the experts selecting that mode answer was then calculated, and the median level of agreement across all six policies became the control value for the comparisons. A value of 100% means that all experts agreed on the same answer choice on the survey question for at least four of the six policies. As the value declines, the level of disagreement on the correct answer choice increases. For the knowledgeable users and crowd workers, the level of agreement with the experts' mode answer was calculated for each question and for each policy. The median levels of agreement for the knowledgeable users and crowd workers on the experts' mode answer across all six policies shows the alignment between these respective groups and the experts. For the knowledgeable users and crowd workers, a value of 100% on a particular question means that all group members agreed on the experts' mode answer choice and shared the same understanding of the privacy policy as the experts for at least four of the six policies.⁸⁵ As the value declines, the group members understand the policies differently from the experts.

For data collection, Table 4 below shows that the three groups shared the same understanding of the collection of health information (100% agreement with experts' answers). With respect to contact information, the knowledgeable users matched the experts (100%), but the crowd workers lagged in their comprehension (90%). The intra-group disagreements on health information reflect that ambiguities in the policies are being interpreted in the same way. The knowledgeable users had a consensus on financial information that matched the experts' most frequently chosen answer choice (100%). Since the experts were not unanimous on the answer choices with respect to financial information, the knowledgeable users may

deletion between groups is as follows: Expert-Knowledgeable $K = 0.736$, Expert-Crowd $K = 0.581$, and Knowledgeable-Crowd $K = 0.555$ for collection practices; Expert-Knowledgeable $K = 0.485$, Expert-Crowd $K = 0.465$, and Knowledgeable-Crowd $K = 0.452$ for sharing practices; and for all pairwise groups, $K = 0.000$ for deletion. The between-group Kappa shows higher levels of agreement above chance for experts and knowledgeable users for collection practices and no distinct differences between groups for sharing practices. For deletion, all raters agreed by assigning the same category to all items which raises the chance probability of assigning the category to 1.000. Consequently, there is no above chance agreement for this category of items.

85. Agreement between knowledgeable users and crowd workers does not necessarily mean the agreed-upon answer is objectively correct. To find out whether these two groups were likely to choose correct answers, we compared their answers to the answers given by the experts on the same policies.

not have been cognizant of the ambiguities seen by the experts. The crowd workers, however, were far off from the experts (40% median level of agreement with the experts' mode answer choice across all policies). This indicates that typical users have a much more difficult time understanding the collection of financial information. Lastly, the knowledgeable users had a significantly different interpretation of the collection of location information than the experts (70% median level of agreement with the experts' mode answer choice across all policies). The crowd workers were, however, in closer agreement with the expert answer choices (90%).

Table 4

Data Collection: Inter-group Agreement

	Collect Contact	Collect Financial	Collect Location	Collect Health
Experts Selecting Mode Answer Choice (Median Across All Policies)	100 %	87.5 %	100 %	100 %
Knowledgeable Users Selecting Experts' Mode Answer Choice (Median Across All Policies)	100 %	100 %	70 %	100 %
Crowd Workers Selecting Experts' Mode Answer Choice (Median Across All Policies)	90 %	40 %	90 %	100 %

For data sharing, Table 5 below demonstrates that the experts had no complete agreement on answer choices for any of the types of data. However, with respect to contact information, the crowd workers agreed with the most frequently chosen expert answer (100% median level of agreement with experts), while the knowledgeable users lagged in their understanding (80%). This may indicate that the knowledgeable users reflected the difficulties that the experts had with policy ambiguity. For financial information, the knowledgeable users had only a modest level of agreement with the experts (50% median level of agreement), while the crowd workers were even more divergent (40%). For location information, the experts' agreement on an answer choice was modest (75% median level of agreement on the mode). Both the knowledgeable users and the crowd workers had a significantly different understanding (60% median level of agreement with experts' mode answer choice). Lastly, for health information, the knowledgeable users and crowd workers converged on the most

frequently chosen expert answer (100% median level of agreement). However, since the experts were divided on the answer choice, this finding suggests that the other groups may have missed nuances in the privacy policies with respect to potential sharing of health information.

Table 5

Data Sharing: Inter-group Agreement

	Sharing Contact	Sharing Financial	Sharing Location	Sharing Health
Experts Selecting Mode Answer Choice (Median Across All Policies)	87.5 %	87.5 %	75 %	75 %
Knowledgeable Users Selecting Experts' Mode Answer Choice (Median Across All Policies)	80 %	50 %	60 %	100 %
Crowd Workers Selecting Experts' Mode Answer (Median Across All Policies)	100 %	40 %	60 %	100 %

For data deletion, Table 6 illustrates that all groups agreed on the understanding of the privacy policies' language (100% median level of agreement). However, since the answer choices with respect to full and partial deletion were collapsed into one, this level of agreement may only reflect that all groups had the same understanding of the existence of, but not the conditions for, data deletion. The intra-group deviation on data deletion means that the policies have important ambiguity on the terms for data deletion, namely whether the policy covers full or partial deletion.

Table 6
Data Deletion: Inter-group Agreement

Experts Selecting Mode Answer Choice (Median Across All Policies)	100 %
Knowledgeable Users Selecting Experts' Mode Answer Choice (Median Across All Policies)	100 %
Crowd Workers Selecting Experts' Mode Answer (Median Across All Policies)	100 %

C. QUALITATIVE DATA ANALYSIS

1. *Difficulty for Survey Respondents*

For each policy, after annotators completed the nine survey questions, they were asked to rate on a 5-point Likert scale the difficulty of annotating that policy, where 1 was “very difficult” and 5 was “very easy.” Averaged over the six policies, knowledgeable users rated the policies as easier to annotate on average (3.23) than the crowd workers (2.53). Interestingly, the experts also found the policies more difficult to annotate (2.88) than the knowledgeable users. Participants’ self-reported ratings concerning their understanding of legal texts varied. The knowledgeable users rated their ability to understand legal texts higher on average (3.8) than the untrained crowd workers (2.29) or the experts (3.5). Looking at the average number of correct answers achieved in the general reading comprehension Cloze test, the experts and knowledgeable users performed on a similar level. Both groups exhibit a median score of seven correct answers (out of eight), with slight variations in the distributions (experts: mean = 6.75, std. dev. = 1.09; knowledgeable users: mean = 6.6, std. dev. = 0.49). The crowd workers exhibit lower reading comprehension with a median score of five correct answers (mean = 4.89, std. dev. = 2.02).

These results suggest that crowd workers may be hampered by less-developed reading comprehension skills in general and, more specifically, for legal text as indicated by their self-assessments. This lower reading comprehension matches the higher perceived difficulty of answering the annotation questions. Because the knowledgeable users and experts exhibit similar general reading comprehension skills, and assuming that experts have more experience in interpreting policy and legal text, the knowledgeable users may have either overestimated their abilities or the experts may have been more cautious in their self-assessment and difficulty ratings.

Table 7 below further shows the individual average difficulty ratings of the different groups for each of the six policies. One set of privacy policies (Lowe's, ABC News, and Washington Post) were perceived as considerably more difficult to annotate by the untrained crowd workers as compared to the other two groups. The Barnes and Noble policy was perceived as difficult to annotate by all three groups. For the Overstock and Weather Underground policies, the difficulty ratings of the experts and untrained crowd workers are quite similar, whereas the knowledgeable users rated them as easier in both cases. Also note that none of the average values reach the "easy" level (4.0).

Table 7

Average difficulty rating of answering the annotation questions for the given policies.

Policy	Experts (avg. ease)	Knowledgeable users (avg. ease)	Crowd workers (avg. ease)
Overstock	3.00	3.40	3.00
Lowe's	3.00	3.00	1.80
Barnes and Noble	2.25	2.60	2.40
ABC News	3.50	3.60	3.00
Weather Underground	2.50	3.20	2.80
Washington Post	3.00	3.60	2.20

2. Trends in Selected Text⁸⁶

a) Consensus on Text Selection

In some instances, the experts achieved either exact or near-exact agreement on the language selected as well as the answer option. These instances of agreement are promising, as they suggest that some scenarios might be used to train natural language processing ("NLP") techniques and thus might be able to provide reliable interpretations of privacy policies. In particular, these instances suggest that crowdsourcing tools might be developed where only (or mostly) relevant text is shown to crowd workers, rather than tools where a crowd worker is required to read an entire policy to answer a particular question.⁸⁷ Where the experts achieved exact agreement,

86. This Section presents trends based on text selected by the experts and the knowledgeable users.

87. See Rohan Ramanath et al., *Identifying Relevant Text Fragments to Help Crowdsourc* *Privacy Policy Annotations*, in PROCEEDINGS OF THE ASS'N FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE CONFERENCE ON HUMAN COMPUTATION & CROWDSOURCING (2014).

they selected the same words or phrases and the same answer option.⁸⁸ Where the experts achieved near-exact agreement, it was typically the case that one annotator selected extra, immaterial words that another did not select, even though the experts selected the same answer option.⁸⁹

Likewise, knowledgeable users were able to achieve exact or near-exact agreement on the language selections as well as the answer option.⁹⁰ As was apparent among experts, it was also typical for one or more knowledgeable users to select additional, immaterial words that others may have not selected, despite the fact that all annotators chose the same answer option.⁹¹ Notably, there were also instances in which annotators would select the same text in support of different answers. This conflict was evident between both expert annotator groups and knowledgeable annotator groups. Ultimately, these occurrences testify to the difficulty of gathering uniform interpretations of privacy policy language, and present a potential challenge to the development of an NLP model that can provide reliable interpretations of privacy policies.⁹² At the same time, these issues suggest that it might also be

88. *See, e.g.*, barnesandnoble.com Privacy Policy *infra* Appendix II. (Appendices for this Article are available at <http://btlj.org/wp-content/uploads/2015/05/privacy-policies-appendix.pdf>. An archived copy is available at <https://archive.org/download/privacy-policies-appendix/privacy-policies-appendix.pdf>.) All annotators selected the following: “we may collect personal information from you, for example your name, e-mail address, billing address, shipping address, phone number” Annotators’ Responses to Survey Question 1 (on file with author).

89. *See, e.g.*, abcnews.go.com Privacy Policy *infra* Appendix I. For Survey Question 7, the annotators selected from the following range of text: “When you allow us to share your personal information with another company, such as: Electing to share your personal information with carefully selected companies so that they can send you offers and promotions about their products and services.” Two annotators selected the darkly shaded language, where only one selected the more lightly shaded language as well. Responses to Survey Question 7 (on file with author).

90. *See, e.g.*, washingtonpost.com Privacy Policy *infra* Appendix V. As commonly seen, sentence selection was identical among all five knowledgeable users. To support their answers, every annotator selected the exact same text, which stated that “Washingtonpost.com asks for information such as your name, e-mail address, year of birth, gender, Zip code, country, street address.” Responses to Survey Question 1 (on file with author).

91. *See, e.g.*, lowes.com Privacy Policy *infra* Appendix III. All knowledgeable users selected the following range of text: “You may choose to provide us with personal information (such as name, contact details and payment information), such as: Contact information, such as your name, address, telephone number, and email address, and your title or occupation.” Responses to Survey Question 1 (on file with author).

92. Both expert and knowledgeable users tended to select the same text to support different answers when asked if a website permitted users to delete personal data. For example, all the experts in response to Survey Question 9 for the washingtonpost.com privacy policy, selected the statement that “[i]f you do not wish to receive the foregoing and therefore unregister from the site, please contact Customer Care and ask to have your

possible to build interfaces that boost the productivity of crowd workers by selectively displaying text fragments that are mostly relevant to the particular questions they are requested to annotate.

b) Interpretation of Policy Silence

Experts differed from knowledgeable users with respect to interpretation of policy silence. Generally, where a policy was silent on a particular practice, experts interpreted such silence to mean that the policy permitted the practice. This is the legal interpretation. Knowledgeable users, on the other hand, often misinterpreted silence to mean that the policy was unclear regarding the practice. Examples of this interpretive difference in the context of data collection, data sharing, and data deletion are described below.

i) Data Collection

One of the best examples of the interpretive difference in the data collection context can be found in Survey Question 2 (“Does the policy state that the website might collect financial information?”) of the wunderground.com policy, which contains no explicit mention of “financial information.”⁹³ In responding to that question, three out of four experts selected sentences that they interpreted to permit the website to collect financial information;⁹⁴ only one expert believed that the policy was unclear

registration account deleted.” While one expert annotator believed that this statement indicated that users could partially remove their data, the remaining three experts selected this statement in support of the answer that it was unclear whether or not users could delete personal data.

With regard to deletion questions, even more frequently than experts, knowledgeable users tended to select the same sentences to support different answers. When responding to Survey Question 9, for example, all annotators for the www.abcnews.go.com privacy policy selected the same sentences which stated that “[y]ou may correct, update and delete your registration account” and “[y]ou may request access to the personal information we hold about you and that we amend or delete it and we request third parties with whom we have shared the information do the same.” Though they all selected the identical text, two annotators believed that these statements indicated “full removal” of personal data, while two other annotators believed that they indicated “partial removal,” and the remaining annotator believed that they indicated that it was “unclear” if users could or could not delete their data. Responses to Survey Question 9 (on file with author).

93. See wunderground.com Privacy Policy *infra* Appendix VI.

94. See wunderground.com (experts) Survey Question 2 (on file with author). Some relevant textual selections chosen by the three experts who said the policy collected financial information are:

“We use information collected on the Services . . . to help fulfill your requests or in connection with the operation of the Services, for example

on the practice. However, knowledgeable users provided almost exactly opposite responses: one knowledgeable user selected sentences that led him or her to think the website could collect financial information,⁹⁵ while three knowledgeable users selected “unclear” and one selected “not applicable.”

ii) Data Sharing

One of the best illustrations of the interpretive difference in the context of data sharing can be found in Question 7 (“Does the policy state that the website might share location information?”) of the Barnes and Noble policy. The policy contains no explicit mention of sharing location information.⁹⁶ In responding to that question, three out of four experts selected sentences that they interpreted to permit the website to share location information;⁹⁷ only

to . . . display information and advertisements that we believe match your interests and profile”

“When your information is collected on the Services, it may be collected directly by or shared with a selected third parties in connection with the operation of the Services or the provision of services to you (“Third Party Processors”). These Third Party Processors may include, for example, companies that operate our online WunderStore, that process credit card information or handle shipping for Weather Underground”

See wunderground.com Privacy Policy *infra* Appendix VI.

95. *See* wunderground.com (knowledgeable users) Survey Question 2. The one annotator selected the following text: “These Third Party Processors may include, for example, companies that operate our online WunderStore, that process credit card information or handle shipping for Weather Underground, that deliver materials to you via e-mail or postal service, that organize, administer, process, or provide advertising services, and/ or that analyze data on our behalf to help us provide more relevant offers to you and to eliminate the delivery of duplicate offers as well as correcting and/or updating users’ data based on information we provide them.”

See wunderground.com Privacy Policy *infra* Appendix VI.

96. *See* barnesandnoble.com Privacy Policy *infra* Appendix II.

97. *See* barnesandnoble.com (experts) Survey Question 7 (on file with author). Two annotators selected Answer Option 3 (Sharing for other purpose) and one chose Answer Option 4 (sharing for other purpose with explicit consent). Some relevant textual selections are:

“These or related applications may also allow you to provide information directly to social networking sites including information about your purchases, physical location.”

“Like many online retailers, we and/or our third party providers use cookies to recognize you as you use or return to the Barnes & Noble Websites.”

“In addition, to provide location-based services on Devices or through Apps, Barnes & Noble and third party application providers may automatically collect real-time geographic location information or other location-based information about you.”

one expert believed that the policy was unclear on the practice. On the other hand, no knowledgeable annotator selected sentences that led them to believe that the website may share users' location information. Instead, two annotators thought the policy was "unclear"⁹⁸ and three selected "not applicable" to reflect the belief that the policy did not appear to address the question.

iii) Data Deletion

There were fewer interpretive differences in the context of data deletion. The annotators frequently agreed on whether a website permitted users to delete personal data.⁹⁹ However, annotators often indicated uncertainty in their responses. They would sometimes respond that it was "unclear" if a website allowed users to delete their information. Other times, annotators answered "partial removal" because it was not clear whether or not websites would retain some of users' data after deletion.¹⁰⁰ These issues arose because some policies addressed editing and correction, but did not explicitly discuss deletion, while other policies stated that users could request deletion, but did not guarantee that the website would actually delete information.

This general doubt among users likely reflects a lack of clarity within the privacy policies themselves. Many policies did not even acknowledge whether users would or would not be able to delete their personal data. For example,

"[Y]our data may be transferred to or shared with a third party as part of a sale, merger, or acquisition of Barnes & Noble or one of its affiliates."

"We provide personal information to our partners that provide product and service offerings or technologies that we think may be of interest to you."

"In connection with purchases of certain Digital Content, we may need to forward information about you (including, for example, your Internet Protocol (IP) address) to the Digital Content provider in order to enable you to download or purchase Digital Content from or through that provider."

See barnesandnoble.com Privacy Policy *infra* Appendix II.

98. *See* barnesandnoble.com (knowledgeable users) Survey Question 7 (on file with author).

99. *See, e.g.*, lowes.com Survey Question 9 (knowledgeable users) (on file with author); wunderground.com Survey Question 9 (knowledgeable users) (on file with author); washingtonpost.com Survey Question 9 (experts) (on file with author).

100. *See, e.g.*, lowes.com (experts and knowledgeable users) Survey Question 9 (on file with author). Notably, when asked if the website permitted users to delete their personal data, two experts answered "unclear," one "partial removal," and one "not applicable." Knowledgeable users expressed similar uncertainty in their answers to this question, with four knowledgeable users answering "unclear" while one answering "not applicable." Survey Question 9 (on file with author).

with respect to the privacy policy for Washington Post, three knowledgeable users selected that it was unclear if personal data was retained, while two answered that the question was not applicable and not addressed by the policy at all.¹⁰¹ Similarly, three out of four experts selected “unclear.”¹⁰² In the washingtonpost.com example, most of the annotators chose the statement “[i]f you do not wish to receive the foregoing and therefore unregister from the site, please contact Customer Care and ask to have your registration account deleted.”¹⁰³ This statement, however, was in reference to deleting a “registration account” in order to stop receiving items from the website such as legal notices and users’ account statuses. This passage was not directed toward the true inquiry of whether or not the website would retain user data for other purposes in the future, and the annotators’ responses reflected this ambiguity.

iv) Assumptions for the Interpretation of Specific Textual Language

Expert and knowledgeable users also made interpretative assumptions about the meaning of policy language. One interpretive difference arose in the context of permissive or conditional language. For example, annotators did not always interpret the word “may” in a policy in the same way. This was illustrated in some cases where experts selected the same text but chose different answer options.¹⁰⁴ This trend applied to both expert and knowledgeable users.¹⁰⁵

101. See washingtonpost.com (knowledgeable users) Survey Question 9 (on file with author).

102. See washingtonpost.com (experts) Survey Question 9 (on file with author).

103. See washingtonpost.com Privacy Policy *infra* Appendix V.

104. See, e.g., wunderground.com Survey Question 2 (on file with author). There, in response to a question about whether the website collects financial information, both annotators selected text that contained the following phrase: “companies that operate our online WunderStore, that process credit card information.” See wunderground.com Privacy Policy *infra* Appendix VI. One annotator, in addition to selecting that phrase, selected the following language preceding it: “These Third Party Processors may include, for example,” and likely interpreted this permissive language as rendering the language unclear (which corresponds to the answer option s/he selected). *Id.* The other annotator, however, did not select this permissive language. Instead, the annotator likely interpreted it to mean that permissiveness means actual practice, and thus chose Answer Option 2 (“Yes—the policy explicitly states that the website might collect financial information”) (on file with author).

105. See, e.g., lowes.com Survey Question 5 (on file with author). Three knowledgeable users selected “unclear” when asked if the privacy policy allowed the website to collect users’ contact information. This is likely due to the permissive language of the privacy policy, which stated that the website “may share personal information we collect on the Site with certain service providers, some of whom may use the information for their own purposes.” See lowes.com Privacy Policy *infra* Appendix III.

Another similar divergence of interpretative assumptions arose with respect to “sharing for core service.” In at least one instance, the experts made different assumptions with respect to the scope of certain language. One annotator believed that third-party services that “operate [the] online WunderStore” or “process credit card information” to support the website fell outside the scope of “core service,” while others did not.¹⁰⁶ Knowledgeable users, too, seemed to differ on their interpretations of what constituted a “core service.” Divergent answer choices for a privacy policy that described third parties as “performing a service” for the website, when the nature of the services were not connected to the user’s interaction with the website, highlighted this confusion.¹⁰⁷

106. *See* wunderground.com Survey Question 6 (on file with author). There, two annotators selected the sentence: “These Third Party Processors may include, for example, companies that operate our online WunderStore, that process credit card information or handle shipping for Weather Underground.” *See* wunderground.com Privacy Policy *infra* Appendix VI. One annotator selected the whole sentence along with Answer Option 3 to notate that “the policy states that the website might share financial information with third parties for other purposes.” *Id.* The other annotator selected only the dark portion with Answer Option 2 to notate that “the policy explicitly states that the website might share financial information with third parties, but only for the purpose of providing a core service.” *Id.*

107. *See* overstock.com Privacy Policy *infra* Appendix IV:

Service Providers

We may share information with companies that provide support services to us (such as a printer, e-mail, mobile marketing, or data enhancement provider) or that help us market our products and services. These companies may need information about you in order to perform their functions. These companies are not authorized to use the information we share with them for any other purpose.

Two out of five knowledgeable users selected the above passage to support their selection that the website shared contact information “for core service only.” All remaining annotators selected the same sentences. However, two of the remaining annotators selected that the website shared user information “for other purposes.” The final remaining annotator selected “unclear.” The answers from the overstock.com survey suggest that two out of five knowledgeable users have been influenced by the phrases “service providers” and “support services” when they selected the option “sharing for core service only.” Though these phrases seem to indicate that sharing would occur for reasons directly related to users’ immediate business with overstock.com, the end of the passage shows that this may not be the case. The final phrase “mobile marketing,” which is listed as a third example of sharing,

The experts also made different assumptions for the interpretation of “unclear” as described in the answer option choices. For example, in one response to the question about whether the website would share contact information with third parties, two experts selected the following:

As part of our ongoing partnership with Microsoft Corporation (“Microsoft”), we may share your personal information with Microsoft and its affiliates and subsidiaries under certain circumstances.¹⁰⁸

However, one of the two experts who selected the text chose Answer Option 3 to reflect that “the policy states that the website might share contact information with third parties for other purposes.” The other expert selecting the same text chose Answer Option 5 (“unclear”), defined as: “the policy does not explicitly state whether or not the website might share contact information with third parties, but the selected sentences could mean that contact information might be shared with third parties.” The second expert likely selected the “unclear” option because of the assumption that “personal information” did not mean or include “contact information.”

The knowledgeable users appear to have similarly diverged in their assumptions for the interpretation of “unclear.” This divergence may particularly be seen in connection with questions about the collection of users’ health information. None of the privacy policies reviewed by the group contained an explicit reference to health information. However, knowledgeable users were often divided on whether or not such collection was possible.¹⁰⁹ For example, in one survey, three knowledgeable users believed that it was “not applicable” to ask whether the privacy policy permitted collection of users’ health information. The description for this answer choice stated that “this question is not addressed by this policy.”¹¹⁰ Yet, two knowledgeable users believed that it was “unclear” whether or not health information was collected. These users selected the statement “[w]e collect the following categories of information,” which was later defined as “[i]nformation you provide in public forums on our sites and applications” and “[i]nformation sent either one-to-one or within a limited group using our message, chat, post or similar functionality.”¹¹¹ These users likely picked

demonstrates that contact information may in fact be shared for advertising purposes.

108. See barnesandnoble.com Privacy Policy *infra* Appendix II; Responses to Survey Question 5 (on file with author).

109. See, e.g., lowes.com Responses to Survey Question 4 (on file with author).

110. See abcnews.go.com Response to Survey Question 5 (on file with author).

111. *Id.*

“unclear” because they read the policy language as ambiguous and as including the possibility that health information would be within the scope of data collection.

Finally, the experts sometimes differed in their interpretations of policy language if inferences could be drawn regarding a website’s practices. For example, with respect to the collection of financial information, one expert selected the choice stipulating “the policy does not explicitly state whether the website might collect financial information or not, [but] the selected sentences could mean that financial information might be collected” and assumed that the selected text meant the website might collect financial information.¹¹² The expert was, in effect, responding on the basis of an assumption about the scope of a term in the policy. As it turned out, the policy contained explicit language about the collection of financial information so the assumption was correct.¹¹³

c) Human Error

Experts appeared to make human errors with occasional instances of mistake. In some cases, experts inadvertently focused on irrelevant material. For example, in response to the question “Does the policy state that the website might collect contact information,” one expert annotator selected explicit language from a policy’s “Types of Information We Collect” section, while another expert annotator focused instead on the policy’s section “WE COLLECT INFORMATION WHEN” and selected language that did not seem to be an appropriate response to the question.¹¹⁴

112. *See* overstock.com Survey Question 2 (on file with author). One expert selected the following pieces of text to come to their conclusion:

“You purchase, order, return, exchange or request information about our products and services from the Sites or mobile applications.”

“Product and Service Fulfillment”

“Fulfill and manage purchases, orders, payments, returns/exchanges, or requests for information, or to otherwise serve you”

See overstock.com Privacy Policy *infra* Appendix IV.

113. *See* overstock.com Survey Question 2 (“What Information We Collect[:] . . . Your credit/debit card number.”).

114. One expert selected the following language from the “Types of Information We Collect” section of overstock.com’s privacy policy:

Your name

Your billing and delivery address

Your e-mail address

Your phone (or mobile) number

See overstock.com Survey Question 1 (on file with author). Another annotator selected the following language from the same policy’s “We Collect Information When” section:

In a few other instances, experts simply made mistakes, as reflected by contradictory choices. As an illustration, one expert chose an answer option that should have been selected only if the accompanying passage mentioned “explicit consent”; the accompanying text passage, however, made no such mention.¹¹⁵ Similarly, in response to a data deletion question, one expert selected an answer choice that should have been chosen if the accompanying selected text mentioned that a user’s data might be retained even after a request that it be deleted; yet, the accompanying selected text did not mention retention.¹¹⁶ It is unclear whether these occurrences resulted from

You purchase, order, return, exchange or request information about our products and services from the Sites or mobile applications.
 You create an Overstock.com account
 You connect with Overstock.com regarding customer service via our customer service center, or on social media platforms.
 You visit the Sites or participate in interactive features of the Sites or mobile applications.
 You use a social media service, for example, Overstock.com's Facebook page or YouTube channel.
 You sign up for e-mails, mobile messages, or social media notifications from Overstock.com.
 You enter a contest or sweepstakes, respond to one of our surveys, or participate in a focus group.
 You provide us with comments, suggestions, or other input

See overstock.com Privacy Policy *infra* Appendix IV.

115. *See* wunderground.com Survey Question 5 (on file with author). Here, one expert annotator selected Answer Option 4, which reads: “Sharing for other purpose (explicit consent)—the policy explicitly states that the website might share contact information with third parties for a purpose that is not a core service, but only if the user provided explicit permission/consent to do so.” The annotator selected the following text to accompany this answer choice:

[W]e may share demographic information, location data, IP address, aggregate (not individual) usage statistics for our Services, other identifiers and information with advertisers and other third parties. For example, we may share IP address, random or anonymous device identifier, city and state, ZIP code, and specific geo-location with the parties identified in subparagraph F below.

See wunderground.com Privacy Policy *infra* Appendix VI. This selection mentions sharing “for a purpose that is not a core service,” but does not require explicit user consent for the sharing it describes.

116. *See* washingtonpost.com Survey Question 9 (on file with author). Here, one expert annotator selected Answer Option 3, which reads: “Partial Removal - the policy explicitly states that users may delete their personal data but some/all of the data might be retained for other purposes, whether the data was provided directly by the user, generated by the user's activities on the website, or acquired from third parties.” That annotator selected the following policy text, which makes no reference to data retention: “If you do not wish to receive the foregoing and therefore unregister from the site, please contact Customer Care and ask to have your registration account deleted. Once your account has been deleted, you

mistaken language selection, mistaken answer option selection, or other factors.

Lastly, experts occasionally missed relevant passages and selected a different portion of text that also accurately responded to the question asked. In these instances, each of the experts could have *also* selected the text that another expert selected.¹¹⁷ This qualifies as a “mistake” because one expert missed relevant language that another saw. This example reveals that privacy policies can be confusing to a degree that even privacy policy “experts” have difficulty recognizing a policy’s full scope. Ultimately, however, these “mistakes” will likely have little effect on adequately informing the NLP tool despite the experts’ differences, as the data used to inform the tool will contain both selections. These examples, though, suggest that crowdsourcing

will no longer have access to washingtonpost.com, however, you may reregister at any time.” See washingtonpost.com Privacy Policy *infra* Appendix V.

117. See, for example, Survey Question 4 regarding collection of health information for the abcnews.com Privacy Policy (on file with author). In response, two experts selected Answer Option 3, which reads: “Unclear—the policy does not explicitly state whether the website might collect health information or not, but the selected sentences could mean that the health information might be collected.” To accompany this selection, one annotator chose one string of text (“Information you provide in public forums on our sites and applications[.] Information sent either one-to-one or within a limited group using our message, chat, post or similar functionality, where we are permitted by law to collect this information”), while the other selected a completely different string (“We acquire information from other trusted sources to update or supplement the information you provided or we collected automatically”). See abcnews.com Privacy Policy *infra* Appendix I. In this instance, both annotators could have selected both strings of text.

Another example can be seen in Survey Question 4 for the lowes.com Privacy Policy. There, in response to a question of whether the website might share contact information, one annotator selected option 3 (“Sharing for other purpose—the policy states that the website might share contact information with third parties for other purposes. The policy makes no statement about the users [sic] consent/permission or user consent is implied.”), and another selected answer option 5 (“Unclear—the policy does not explicitly state whether the website might share contact information with third parties or not, but the selected sentences could mean that contact information might be shared with third parties.”). Again, here, the annotators selected different language supporting each answer choice. One annotator (the one who chose option 5) selected, among others, the following string of text: “We may share personal information we collect on the Site with certain service providers, some of whom may use the information for their own purposes.” See lowes.com Privacy Policy *infra* Appendix III. The other annotator (choosing answer option 3) selected this text: “We reserve the right to transfer personal information we have about you in the event we sell or transfer all or a portion of our business or assets (including, without limitation, in the event of a reorganization, dissolution or liquidation). Should such a sale or transfer occur, we will use reasonable efforts to direct the transferee to use personal information you have provided to us in a manner that is consistent with our Privacy Statement.” See lowes.com Privacy Policy *infra* Appendix III. Again, each annotator could have selected the text that the other selected.

solutions may help remedy some of the human errors made while interpreting policies.

V. SIGNIFICANCE OF FINDINGS

Discrepancies between privacy experts and law and policy researchers reveal areas for careful attention to the quality of privacy policies. Discrepancies between privacy experts and non-expert users cast doubt on whether website notices, as they are typically worded today, can effectively convey privacy policies to the general public. If websites are not effectively conveying privacy policies to consumers in a way that a “reasonable person” could understand, notice and choice fails as a framework. If consumers cannot successfully decode privacy policies, the underpinnings of the U.S. approach to privacy are unsustainable, and regulation may be necessary. Indeed, a gap in interpretation between expert and typical users indicates that privacy policies are in fact misleading the general public and that those policies could be considered legally unfair and deceptive if the gap is intentional.

This Section will address some of the implications from the findings for common understandings of website privacy policies and for crowdsourcing the interpretation of privacy policies.

A. IMPLICATIONS FOR COMMON UNDERSTANDING AND CONSUMER DECEPTION

The findings show a number of areas where website privacy policies are too ambiguous to be meaningful and reveal a need to clarify specific data practices. The research demonstrates that policies describe websites’ data sharing practices poorly. Experts could not reach consensus on interpretation of data sharing practices generally and agreed *even less* as to the various nuances of data sharing.¹¹⁸ Website owners must be more candid and clear in drafting notices of data sharing practices. In particular, more precision is needed with respect to the collection of specific types of users’ personal information. The findings showed common understandings among experts on contact and location information, but not on financial and health information.¹¹⁹ This indicates that more clarity is necessary to spell out websites’ specific practices with respect to sensitive information (i.e., financial and health information). General statements concerning “personal

118. *See supra* Table 2.

119. *See supra* Table 1.

information” often introduce ambiguity into a policy and makes it difficult to interpret which information the website is actually collecting.

The findings also showed many instances where a majority of non-expert users interpreted terms differently from experts.¹²⁰ This implies that website notices are not conveying accurate information to consumers and that privacy policies may be misleading the public in specific areas. For example, knowledgeable users and crowd workers both lagged substantially behind the experts in their understanding of websites’ data practices.¹²¹ Conversely, there were instances when experts could not agree on an interpretation, but non-experts users did agree on the interpretation. This may indicate that some language in website notices are commonly misunderstood where non-expert readers fail to recognize the ambiguity in a site’s stated practices. For instance, when policies are silent on specific issues, or when conditions of sharing with third parties are not described clearly, non-experts will have a tendency to misunderstand the terms.

Lastly, since the findings reveal that complete agreement was uncommon, even among experts, website policies may be difficult to interpret through automated means. In other words, where users are unable to understand the policies accurately, automated tools will similarly misapprehend policy language. These difficulties are most pronounced at the level of specific practices. In instances where there is consensus on how data is treated *broadly*, for example, whether a policy states that a user may delete his or her personal information generally,¹²² confusion increases exponentially when more nuance exists as to the data practice—such as instances when a policy provides for the ability to *fully* or *partially* delete user information.¹²³ This difficulty applied to each group of annotators. This disparity is very significant because personal privacy preferences are contextually based.¹²⁴ To have contextual integrity, the granular aspects of a data practice, not just whether a website collects, shares, or deletes personal information in general, will need to be understandable to a user. Indeed, a policy statement acknowledging general data collection and/or sharing may do very little to inform readers about the practices relevant for the user.

120. This arises when experts reach a *unanimous* consensus, yet a majority of non-experts interpret the same terms differently.

121. *See supra* Table 5.

122. *See supra* Table 3.

123. *See supra* Table 3.

124. *See generally* HELEN NISSENBAUM, *PRIVACY IN CONTEXT: TECHNOLOGY, POLICY AND THE INTEGRITY OF SOCIAL LIFE* (2009).

The lack of agreement among users and the difficulties they have interpreting policies suggest that consumers are currently misled by website privacy policies. To the extent that vague and misleading terminology is the result of drafting errors and omissions, this research shows areas where website policies can and need to be improved. To the extent, however, that vague and misleading terms are intentionally introduced into website privacy policies, this research suggests that websites are successfully deceiving consumers. The Federal Trade Commission under its unfair and deceptive practice jurisdiction, and private litigants under state unfair and deceptive practice laws, may be able to address such deficiencies in enforcement proceedings.¹²⁵

B. IMPLICATIONS FOR CROWDSOURCING

Semi-automated extraction of meaning from website privacy policies may, in some circumstances, be a solution to help solve the difficulties users face in the interpretation and comprehension of privacy policies. Crowdsourcing is a critical part of such extraction and the findings of this research raise a number of implications for crowdsourcing opportunities. These implications arise from cases in our study when experts agree on the interpretation of terms in the privacy policies and when experts disagreed on interpretation.

1. *When Experts Agree*

Where there is intra-group agreement among experts, then it is possible to compare the crowdsourced majority answer to the expert-selected answer to see whether a majority of crowd workers are able to arrive at the “correct” answer. Crowd workers in our study did a reasonable job predicting the answer when the experts had intra-group agreement. For example, experts had a high level of intra-group agreement with respect to the collection of location information¹²⁶ and, at the same time, the majority of crowd workers chose the same response as the experts, signifying that both groups interpreted the collection of location information in the same way.¹²⁷ This alignment between experts and crowd workers on the interpretation suggests that crowdsourcing could be leveraged to identify collection practices for certain types of data in privacy policies, i.e., those in our study where

125. For a discussion of these types of enforcement actions, see PRIVACY ENFORCEMENT ACTIONS, *supra* note 71.

126. *See supra* Table 1.

127. *See supra* Table 4.

agreement between experts and crowd workers indicates that the two groups' interpretations are very close.

By contrast, where there is intra-group agreement among experts and crowd workers fail to arrive at the experts' "correct" answer, then crowd workers misunderstand policy terms. For example, experts were largely in agreement on whether websites collected financial information,¹²⁸ but the majority of crowd workers failed to select the experts' answer.¹²⁹ This inconsistency suggests that extracting information practices for these types of data cannot easily be crowd sourced to typical users because their interpretation of the policy is likely to be wrong compared to the experts' interpretation.¹³⁰

2. *When Experts Disagree*

Crowd workers, also, show positive signs at predicting expert disagreement. For example, neither experts nor crowd workers could agree on whether websites shared location information.¹³¹ This suggests that crowd worker disagreement might serve as a proxy for expert disagreements and that those disagreements can be used to identify aspects of, and potentially specific passages within, websites' privacy policies that lack adequate precision.

By contrast, when experts disagree but the majority of crowd workers do choose an answer, the crowd workers may either be misunderstanding the policy or interpreting the survey questions differently from the experts. For example, experts had low intra-group agreement on whether websites shared health information¹³² and did not agree completely on the best answer choice.¹³³ Yet, crowd workers all converged on the same response to the

128. *See supra* Table 1.

129. *See supra* Table 4.

130. At the same time, one natural language processing technique might be used to recognize text patterns in the policy that prompted experts' answer choices and then build crowdsourcing tools that combine machine learning and natural language processing techniques to improve the performance of crowd workers. The results of an automated extraction would be shown to crowd workers, for whom it may then be easier to verify the correctness of an extracted data practice or be less influenced by other text as compared to identifying the data practice correctly without assistance. The possibility remains to be studied. Another possible opportunity for enhancing the effectiveness of crowd workers and the accuracy of their extractions is to train them with example annotations provided by experts or knowledgeable users.

131. *See supra* Table 2. Experts were divided on the sharing of location information with a median level of agreement on all answer choices at 62.5%. Crowd workers had a median level of agreement at 40%.

132. *See supra* Table 2.

133. *See supra* Table 5.

sharing question.¹³⁴ In a case like this, there appear to be two possible explanations. First, the crowd workers might be failing to see the ambiguity in the policy and consequently would be misunderstanding the terms. Alternatively, as reflected in the qualitative findings, the two groups might be interpreting the questions and answer choices differently.¹³⁵

These issues require further exploration to inform whether and how crowdsourcing tasks might be organized in support of semi-automated privacy policy annotation to achieve replication of expert interpretations.¹³⁶ Similarly, further exploration will be necessary to identify whether and how to augment crowdsourcing with combinations of NLP techniques and machine learning. Our study strongly suggests, however, that for a narrow range of data practices, both crowdsourcing and automated extraction may be within reach. For other data practices, such as information sharing with third parties, where policies are sufficiently ambiguous to cause disagreement between experts, further research will be required to address the challenges of devising crowdsourcing tasks and automated extraction that would enable non-experts to accurately determine a policy's interpretation.

VI. CONCLUSIONS

The results of this study have significant implications for public policy and technological developments. The disagreements among experts as reflected in the intra-group findings¹³⁷ demonstrate that privacy policies are ambiguous on key terms. Specifically, the findings show interpretative challenges with respect to certain types of information, the scope of data sharing, and the scope of data deletion rights. The findings also show that both knowledgeable users and crowd workers had greater difficulty deciphering privacy policy language than the experts.¹³⁸ Taken together, these findings suggest that privacy policies are written ambiguously and in a way that leads both knowledgeable users and crowd workers to misapprehend websites' data practices as well as cause disagreement among experts with respect to certain data practices.

If the ambiguity in website privacy policies is unintentional, then the findings illustrate where businesses need to improve the clarity of their website privacy policies. The study methodology may also enable industry-

134. *See supra* Table 5.

135. *See* Subsection IV.C.2.a)

136. For an experimental approach towards refining a workflow of crowdsourcing tasks for extracting data practices from privacy policies, see Breau & Schabu, *supra* note 62.

137. *See supra* Sections IV.A, C.2.

138. *See supra* Sections IV.C.1.

wide tracking of how business sectors adjust their policies to more clearly explain their information practices. Specifically, web sweeps on a periodic basis to collect privacy policies as discussed in Section III.B would provide a data set to examine the evolution of privacy policies by industry sectors over time. Then, to the extent that some data practices can be extracted from crowdsourcing, and eventually natural language processing, these techniques could identify if privacy policy terms change over time.

If, however, the ambiguity is intentional, then the study findings suggest that website policies deceive the public. Such deception would be ripe for investigation by the Federal Trade Commission as “unfair and deceptive” business practices.

For the development of automated and crowdsourcing technological tools to assist end-users and policymakers in understanding privacy policies, the findings show initial promise as well as a need for further research into the important, identified challenges.¹³⁹ Machine learning and natural language processing techniques might be useful to highlight potentially relevant passages in a privacy policy for crowd workers who are trying to extract information about that practice. Our findings further suggest that crowdsourcing can be used today to infer some aspects of the meaning of textual statements and that these interpretations may be used to code for the automatic extraction of answers from those passages. Where crowd workers can predict expert disagreement about the interpretation of policy language, those ambiguous passages might be parsed out. For the remaining policy text, crowd workers’ success in predicting answers where experts would agree may be used to code passages where there is interpretative agreement. This approach may thus provide a more accurate understanding of the legal meaning of some of the privacy policy terms where an individual, non-expert user would otherwise be misled. However, given a high level of actual disagreement among experts or crowd-predicted disagreement among experts, significant and important terms in privacy policies will very likely escape effective interpretation by crowd workers because the effectiveness of automated tools and crowdsourced interpretation depends on an accurate baseline meaning for text in a privacy policy. Such baseline meaning is currently missing for some of the key attributes of the website policies that were analyzed in this study.

139. See Sadeh et al., *supra* note 2 (describing the importance of this research for natural language processing and crowdsourcing).

APPENDIX I:

PRIVACY POLICIES*

I. ABCNEWS.GO.COM

The Walt Disney Company has a rich tradition of bringing great stories, characters and experiences to our guests around the world, and our sites and applications are created to entertain and connect guests with the best that we have to offer on the platforms and devices our guests prefer.

Our privacy policy is designed to provide transparency into our privacy practices and principles, in a format that our guests can navigate, read and understand. We are dedicated to treating your personal information with care and respect.

Privacy Policy

TRUSTe online privacy certification
Last Modified: December 30, 2013

This privacy policy describes the treatment of information provided or collected on the sites where this privacy policy is posted. It also explains the treatment of information provided or collected on applications we make available on third-party sites or platforms if disclosed to you in connection with use of the application. We follow this privacy policy in accordance with local law in the places where we operate.

Types of Information We Collect

How We Collect Your Information

Use of Your Information by The Walt Disney Family of Companies

Sharing Your Information with Other Companies

Your Controls and Choices

* These privacy policies are the versions used in this study as collected from the websites of companies. The policies in this appendix do not have the same formatting as the originals found on the websites. The table of contents, the alphabetical organization and some of the heading fonts were added to allow for faster navigation. No changes were made to the text of any of the policies.

Children's Privacy
Data Security and Integrity
Data Transfers, Storage and Processing Globally
Changes to this Privacy Policy
Comments and Questions

TYPES OF INFORMATION WE COLLECT

We collect two basic types of information - personal information and anonymous information - and we may use personal and anonymous information to create a third type of information, aggregate information. We collect the following categories of information:

Registration information you provide when you create an account, including your first name and surname, country of residence, gender, date of birth, email address, username and password

Transaction information you provide when you request information or purchase a product or service from us, whether on our sites or through our applications, including your postal address, telephone number and payment information

Information you provide in public forums on our sites and applications
Information sent either one-to-one or within a limited group using our message, chat, post or similar functionality, where we are permitted by law to collect this information

Information you provide to us when you use our sites and applications, our applications on third-party sites or platforms such as social networking sites, or link your profile on a third-party site or platform with your registration account

Location information when you visit our sites or use our applications, including location information either provided by a mobile device interacting with one of our sites or applications, or associated with your IP address, where we are permitted by law to process this information

Usage, viewing and technical data, including your device identifier or IP address, when you visit our sites, use our applications on third-party sites or platforms or open emails we send

[back to top of page](#)

HOW WE COLLECT YOUR INFORMATION

We collect information you provide to us when you request products,

services or information from us, register with us, participate in public forums or other activities on our sites and applications, respond to customer surveys, or otherwise interact with us. Please keep in mind that when you provide information to us on a third-party site or platform (for example, via our applications), the information you provide may be separately collected by the third-party site or platform. The information we collect is covered by this privacy policy and the information the third-party site or platform collects is subject to the third-party site or platform's privacy practices. Privacy choices you have made on the third-party site or platform will not apply to our use of the information we have collected directly through our applications.

We collect information through technology, such as cookies, Flash cookies and Web beacons, including when you visit our sites and applications or use our applications on third-party sites or platforms. Please visit [Online Tracking and Advertising](#) for further information, including [Do Not Track](#) and how to disable cookies.

We acquire information from other trusted sources to update or supplement the information you provided or we collected automatically. Local law may require that you authorize the third party to share your information with us before we can acquire it.

[back to top of page](#)

USE OF YOUR INFORMATION BY THE WALT DISNEY FAMILY OF COMPANIES

A member of The Walt Disney Family of Companies, which includes many different brands, will be the data controller for your information. The relevant data controller(s) can be determined here. Other members of The Walt Disney Family of Companies may have access to your information where they perform services on behalf of the data controller(s) (as a data processor) and, unless prohibited under applicable law, for use on their own behalf (as a data controller) for the following purposes:

Provide you with the products and services you request

Communicate with you about your account or transactions with us and send you information about features on our sites and applications or changes to our policies

Consistent with local law and choices and controls that may be available to you:

Send you offers and promotions for our products and services or third-party products and services

Personalize content and experiences on our sites and applications

Provide you with advertising based on your activity on our sites and applications and on third-party sites and applications. To learn more about

how we use your information for personalization and tracking, please visit [Online Tracking and Advertising](#).

[Optimize or improve our products, services and operations](#)

[Detect, investigate and prevent activities that may violate our policies or be illegal](#)

[back to top of page](#)

SHARING YOUR INFORMATION WITH OTHER COMPANIES

We will not share your personal information outside The Walt Disney Family of Companies except in limited circumstances, including:

When you allow us to share your personal information with another company, such as:

[Electing to share your personal information with carefully selected companies so that they can send you offers and promotions about their products and services](#)

[Directing us to share your personal information with third-party sites or platforms, such as social networking sites](#)

Please note that once we share your personal information with another company, the information received by the other company becomes subject to the other company's privacy practices.

When we cooperate with financial institutions to offer co-branded products or services to you, such as our co-branded Disney Rewards Visa Card; however, we will do so only if permitted by applicable law and, in these cases, the financial institutions are prohibited from using your personal information for purposes other than those related to the co-branded products or services. When companies perform services on our behalf, like package delivery and customer service; however, these companies are prohibited from using your personal information for purposes other than those requested by us or required by law.

When we share personal information with third parties in connection with the sale of a business, to enforce our Terms of Use or rules, to ensure the safety and security of our guests and third parties, to protect our rights and property and the rights and property of our guests and third parties, to comply with legal process or in other cases if we believe in good faith that disclosure is required by law.

[back to top of page](#)

YOUR CONTROLS AND CHOICES

We provide you the ability to exercise certain controls and choices regarding our collection, use and sharing of your information. In accordance with local law, your controls and choices may include:

You may correct, update and delete your registration account

You may change your choices for subscriptions, newsletters and alerts

You may choose whether to receive from us offers and promotions for our products and services, or products and services that we think may be of interest to you

You may choose whether we share your personal information with other companies so they can send you offers and promotions about their products and services

You may choose whether to receive targeted advertising from many ad networks, data exchanges, marketing analytics and other service providers here

You may request access to the personal information we hold about you and that we amend or delete it and we request third parties with whom we have shared the information do the same

You may exercise your controls and choices, or request access to your personal information, by visiting Communication Choices, contacting Guest Services, or following instructions provided in communications sent to you. Please be aware that, if you do not allow us to collect personal information from you, we may not be able to deliver certain products and services to you, and some of our services may not be able to take account of your interests and preferences. If you have questions regarding the specific personal information about you that we process or retain, please contact Guest Services.

[back to top of page](#)

CHILDREN'S PRIVACY

We recognize the need to provide further privacy protections with respect to personal information we may collect from children on our sites and applications. Some of the features on our sites and applications are age-gated so that they are not available for use by children, and we do not knowingly collect personal information from children in connection with those features. When we intend to collect personal information from children, we take additional steps to protect children's privacy, including:

Notifying parents about our information practices with regard to children, including the types of personal information we may collect from children, the uses to which we may put that information, and whether and with whom we may share that information

In accordance with applicable law, obtaining consent from parents for the collection of personal information from their children, or for sending information about our products and services directly to their children

Limiting our collection of personal information from children to no more than is reasonably necessary to participate in an online activity

Giving parents access or the ability to request access to personal information we have collected from their children and the ability to request that the personal information be changed or deleted

For additional information about our practices in the United States and Latin America regarding children's personal information, please read our Children's Privacy Policy.

[back to top of page](#)

DATA SECURITY AND INTEGRITY

The security, integrity and confidentiality of your information are extremely important to us. We have implemented technical, administrative and physical security measures that are designed to protect guest information from unauthorized access, disclosure, use and modification. From time to time, we review our security procedures to consider appropriate new technology and methods. Please be aware though that, despite our best efforts, no security measures are perfect or impenetrable.

[back to top of page](#)

DATA TRANSFERS, STORAGE AND PROCESSING GLOBALLY

We operate globally and may transfer your personal information to individual companies of The Walt Disney Family of Companies or third parties in locations around the world for the purposes described in this privacy policy. Wherever your personal information is transferred, stored or processed by us, we will take reasonable steps to safeguard the privacy of your personal information. Additionally, when using or disclosing personal information transferred from the European Union, we abide by the Safe Harbor Principles as set forth by the U.S. Department of Commerce, use standard contract clauses approved by the European Commission, adopt other means

under European Union law for ensuring adequate safeguards, or obtain your consent. We also apply the substantive requirements of the Safe Harbor Principles when transferring personal information from Australia.

[back to top of page](#)

CHANGES TO THIS PRIVACY POLICY

From time to time, we may change this privacy policy to accommodate new technologies, industry practices, regulatory requirements or for other purposes. We will provide notice to you if these changes are material and, where required by applicable law, we will obtain your consent.

[back to top of page](#)

COMMENTS AND QUESTIONS

If you have a comment or question about this privacy policy, please contact Guest Services. Our sites and applications may contain links to other sites not owned or controlled by us and we are not responsible for the privacy practices of those sites. We encourage you to be aware when you leave our sites or applications and to read the privacy policies of other sites that may collect your personal information.

Notice to California Residents: If you are a California resident, California Civil Code Section 1798.83 permits you to request information regarding the disclosure of your personal information by certain members of The Walt Disney Family of Companies to third parties for the third parties' direct marketing purposes. With respect to these entities, this privacy policy applies only to their activities within the State of California. To make such a request, please send an email to caprivacy.wdig@twdc.com or write us:

CA Privacy Rights
Disney Interactive
500 South Buena Vista Street
Mail Code 7667
Burbank, CA 91521-7667

In your request, please specify the member of The Walt Disney Family of companies to which your request pertains. If no member is specified, we will treat your request as pertaining to Disney Online.

[back to top of page](#)

DEFINITIONS

Aggregate Information. Aggregate information means information about groups or categories of guests, which does not identify and cannot reasonably be used to identify an individual guest.

[back to top of page](#)

Anonymous Information. Anonymous information means information that does not directly or indirectly identify, and cannot reasonably be used to identify, an individual guest.

[back to top of page](#)

Application. Application means a program or service operated by us (or on our behalf) that may be displayed on various online, mobile or other platforms and environments, including those operated by third parties, which permits us to interact directly with our guests.

[back to top of page](#)

Children. Children means individuals who we have identified are not of legal age to consent to the collection and processing of their personal information. In the United States and Latin America, the term "children" refers to individuals under 13 years of age.

[back to top of page](#)

Data Controller. The data controller is the subsidiary or affiliated entity of The Walt Disney Company that is responsible for the personal information collected from sites and applications, as follows:

Sites and Applications Company Contact Information

Disney Club Penguin ("Club Penguin") Disney Canada Inc. (formerly known as Disney Online Studios Canada Inc.) Club Penguin
c/o Disney Online Studios Canada Inc.

1628 Dickson Avenue, Suite 500

Kelowna, British Columbia V1Y 9X1

CANADA

support@clubpenguin.com

Disney Movies Online Disney Online, Buena Vista Home Entertainment,
Inc. Disney Interactive

500 South Buena Vista Street

Mail Code 7667
Burbank, CA 91521-7667
United States of America
Guest Services
Disney Studio All Access
Disney Movie Rewards Disney Online, Buena Vista Home Entertainment,
Inc., Walt Disney Studios Motion Pictures Disney Interactive
500 South Buena Vista Street

Mail Code 7667
Burbank, CA 91521-7667
United States of America
Guest Services

All other sites and applications Disney Online Disney Interactive
500 South Buena Vista Street
Mail Code 7667

Burbank, CA 91521-7667
United States of America
Guest Services

[back to top of page](#)

Data Processor. A data processor is a person or entity that processes personal information on behalf of a data controller (or data controllers) and is permitted to perform data processing only as directed by the data controller(s).

[back to top of page](#)

IP Address. An IP address is associated with the access point through which you enter the Internet, and is typically controlled by your Internet Service Provider (ISP), your company, or your university. We may use IP addresses to collect information regarding the frequency with which our guests visit various parts of our sites and applications, and we may combine IP addresses with personal information.

[back to top of page](#)

Member. Member means a subsidiary or affiliated entity that is part of The Walt Disney Family of Companies.

[back to top of page](#)

Notice. Notice may be by email to you at the last email address you provided us, by posting notice of such changes on our sites and applications, or by other means, consistent with applicable law.

[back to top of page](#)

Parents. Parents means a parent or legal guardian.

[back to top of page](#)

Personal information. Personal information means information that identifies (whether directly or indirectly) a particular individual, such as the individual's name, postal address, email address and telephone number. When anonymous information is directly or indirectly associated with personal information, this anonymous information also is treated as personal information.

[back to top of page](#)

Public Forums. Our sites and applications may offer message boards, conversation pages, blogs, chat rooms, social community environments, profile pages, and other forums that do not have a restricted audience. If you provide personal information when you use any of these features, that personal information may be publicly posted and otherwise disclosed without limitation as to its use by us or by a third party. To request removal of your personal information from a public forum on one of our sites or applications, please contact Guest Services.

[back to top of page](#)

The Walt Disney Family of Companies. The Walt Disney Family of Companies refers to The Walt Disney Company and its subsidiary and affiliated entities, which offer their products and services under various brand names. These companies engage in a number of businesses, including theme parks and travel, motion pictures and television, publishing, consumer products and interactive services. The Walt Disney Company brands include, among others, the following:

ABC

Babble

Baby Einstein

BabyZone

Club Penguin

Disney

Disney Pixar

ESPN

Hollywood Records

Indiana Jones

Lucasfilm

Marvel
Muppets
Playdom
Spoonful
Star Wars
Tapulous
Touchstone
back to top of page

GUEST SERVICES CONTACT INFORMATION

United States of America:

Guest Services
Disney Interactive
500 South Buena Vista Street
Mail Code 7667
Burbank, CA 91521-7667
United States of America
Guest Services

For questions related to children's privacy, you may also telephone Guest Services at (877) 466-6669.

Disney Interactive has received the TRUSTe Privacy Seal in the United States, signifying that this privacy policy and our privacy practices have been reviewed for compliance with the TRUSTe program listed on the validation page available when you click on the TRUSTe Privacy Seal. The TRUSTe program covers only those properties identified on the validation page. If you believe that Disney Interactive has not responded to your inquiry or your inquiry has not been satisfactorily addressed, you may contact TRUSTe here or the United States Federal Trade Commission through its online consumer complaint form available here

II. BARNESANDNOBLE.COM

Effective Date: May 7, 2013

This Privacy Policy applies to personal and other information that may be collected when you interact with the Barnes & Noble enterprise which consists of (a) Barnes & Noble, Inc. and its subsidiaries, including their

respective businesses and operations (collectively, "Barnes & Noble"); (b) businesses and operations managed or operated by Barnes & Noble; (c) websites owned, operated and managed by Barnes & Noble, including each website that links to this Privacy Policy, and any digital content stores operated by Barnes & Noble (collectively, "Barnes & Noble Websites"); and (d) Barnes & Noble's devices, content and applications. Protecting the privacy and security of your personal information is a priority at Barnes & Noble, and we believe that a single, comprehensive privacy policy that is straightforward and clear is in the best interests of our customers and our businesses. More detailed information about specific practices regarding Barnes & Noble and NOOK mobile devices ("Devices") and mobile applications developed and maintained by Barnes & Noble and NOOK ("Apps") is contained in the Barnes & Noble Mobile Privacy Supplement which is incorporated into, and subject to, this Privacy Policy. To review the prior version of Barnes & Noble's Privacy Policy, please [click here](#).

By doing business with or interacting with Barnes & Noble in the manner described in this Privacy Policy at any time, you are accepting the practices described in this Privacy Policy and you consent to the application of this Privacy Policy to the collection, storage, use and disclosure of all your personal and other information as described.

Special Note: On October 4, 2012, Barnes & Noble underwent a corporate reorganization resulting in its digital and College businesses becoming owned by Barnes & Noble's majority owned subsidiary NOOK Media LLC and its subsidiaries (collectively "NOOK Media"). These current privacy policies and the privacy practices described herein apply to personal information previously collected from you by Barnes & Noble, Inc. and each of its subsidiaries, including NOOK Media, as well as personal information that they may collect from you in the future. Please be advised that NOOK Media intends to launch one or more of its own websites in the future for purposes of operating its digital and College businesses. NOOK Media expects to post privacy policies applicable to those websites when they are launched.

Barnes & Noble reserves the right to modify or amend this Privacy Policy at any time, but you can be assured that, should it be necessary to do so, we will always do so in accordance with the Barnes & Noble Privacy Principles of Clarity, Security and Integrity. You will be notified of any material changes to this Privacy Policy which are less protective of customer information prior to such changes becoming effective. We may notify you of these changes by

email reminders, by notice on this site, or by other acceptable means. We encourage you to periodically review this page for the latest information on our privacy practices.

Clarity: We strive to communicate clearly about your privacy and how we handle your personal information.

Security: We follow security standards, processes and procedures that are designed to protect your personal information.

Integrity: We do not sell or rent your personal information and respect your preferences with respect to your personal information.

Below you will find answers to the following frequently asked questions about how we collect, use and share your personal information:

1. What is the personal information that we collect?
2. Why do we collect personal information?
3. How do we collect your personal information?
4. How do we use personal information?
5. With whom do we share personal information?
6. How do we secure personal information?
7. How do we respect your choices about your personal information?
8. Whom do I contact if I have questions or concerns?
9. What other information about Barnes & Noble's Privacy Policy would I want to know?

1. WHAT IS THE PERSONAL INFORMATION THAT WE COLLECT? *

Depending on how you choose to interact with the Barnes & Noble enterprise, we may collect personal information from you, for example your name, e-mail address, billing address, shipping address, phone number, credit card information, date of birth and other persistent identifiers that can be used to personally identify you. If you rent, purchase or otherwise place orders for textbooks from Barnes & Noble College Booksellers, we may also collect your student or faculty identification number, financial aid number, your driver license number and information regarding courses you enroll in or teach. We may also collect other personal information regarding your interaction with Barnes & Noble Websites, Devices and Apps, including

* Numbers have been left in the Barnes & Noble privacy policy to reflect its original formatting, separate from BTLJ formatting.

usage information, the items that you browse, purchase, download, read, watch and otherwise access, and your geographic location. For further details about personal information collected through Apps and Devices, please read the Barnes & Noble Mobile Privacy Supplement.

2. WHY DO WE COLLECT PERSONAL INFORMATION?

We collect your personal information in an effort to provide you with a superior customer experience and, as necessary, to administer our business. It allows us to provide you with easy access to our products and services, with a particular focus on the items and programs that may be of most interest to you. For more information as to why we collect personal information, please see the section below entitled: "How do we use your personal information?"

Your personal information also allows us to communicate with you about special offers, promotions, and other marketing programs and news that may be of interest to you. You always have the opportunity to unsubscribe from promotional emails by following the instructions included in each marketing email or by changing your preferences in your account.

3. HOW DO WE COLLECT YOUR PERSONAL INFORMATION?

(a) Information that you provide to us

As a general matter, you can browse in our stores and on the Barnes & Noble Websites without submitting your personal information to us, although we may receive and collect certain personal information automatically, as outlined in Section 3(b) of this Privacy Policy, including site analytics regarding our websites, information your Internet browser automatically sends when you visit our websites, and information collected by cookies.

Making a purchase or placing an order
Creating an account or joining the Member Loyalty Program
Applying for a Barnes & Noble MasterCard
Signing up for Invite-a-Friend Emails
Using specific features of your Device
Contacting customer service
Additional Barnes & Noble Website, Device or App Features
Interacting with social networking sites
Entering a sweepstakes or contest

For your convenience, we have provided a summary description of each of these circumstances below.

Making a purchase or placing an order

When you make a purchase, rent or place an order in our stores, place an order online or purchase, download, rent or stream books, periodicals, movies, television shows and other digital content ("Digital Content"), you may need to submit personal information to us.

Creating an account or joining the Member Loyalty Program

In order to use Devices, certain Apps and features of Barnes & Noble Websites, or to purchase or access Digital Content, you may be required to create a password-protected user account and provide us with personal information when you do so. Similarly, when you enroll in our Member Loyalty Program, we will ask you to submit personal information as part of your member profile. We will store and use this information to administer the programs and services in which you choose to participate, and as permitted by this Privacy Policy.

Applying for a Barnes & Noble MasterCard

When you apply for a Barnes & Noble MasterCard, you will need to submit personal information directly to the issuer of the MasterCard. Personal information that you provide directly to the issuer is subject to the bank's privacy policies, practices and procedures, not Barnes & Noble's.

In addition, we disclose certain personal information to the issuer of the MasterCard in connection with the administration of the Barnes & Noble MasterCard program. The issuer does not have the right to use the personal information we provide beyond what is necessary to assist us or to administer this program. The issuer is contractually obliged to maintain the confidentiality and security of the personal information we provide and is restricted from using such information in any way not expressly authorized by us.

Signing up for Invite-a-Friend Features

Our Invite-a-Friend feature is located in various areas of certain Barnes & Noble Websites and within certain Devices. By clicking on 'Email' or 'Find friends from contacts' on your Device, customers can email a link to a friend which will invite the friend to view the Barnes & Noble Website page or attend the event described on that page.

The 'Find friends from contacts' feature on the Device may access the contacts you have entered into your Device in order to generate such emails. To remove yourself from these mailings, please call our Customer Service Center at 1-800-THE-BOOK (1-800-824-2665). Individuals calling from abroad should call us at 1-201-559-3882.

Using specific features of Barnes & Noble Websites, Devices and Apps

Barnes & Noble Websites, Devices and Apps may provide you with the ability to enter (either directly, or by authorizing Barnes & Noble to download the information from a third party such as a social networking website) your information such as your contacts, calendar entries or photos, or images you may submit to us to help us locate products for you. For further details about personal information collected through Devices and Apps, please read the Barnes & Noble Mobile Privacy Supplement.

Contacting Customer Service

When you contact customer service, we may ask you to provide, or confirm, personal information so that we can better serve you.

Additional Website, Device or App Features

From time to time we may offer or provide new or additional Barnes & Noble Website, Device or App features and functionality. Such features or functionality may request or require you to submit personal information to us in order to use such feature or functionality.

Entering a Sweepstakes or Contest

If you enter a sweepstakes or contest we offer, we may ask you to provide personal information so that we can consider your entry and, if you win, so that you may redeem your prize.

Interacting with Social Networking Sites

Our Devices and Apps may provide you with the ability to enter (directly, or by authorizing us to download the information from a third party such as a social networking site or application) personal information such as contacts or lists of friends. These or related applications may also allow you to provide information directly to social networking sites including information about your purchases, physical location, or comments. However, we will not provide any of your information to social networking sites without your express consent.

(b) Information automatically collected

There are circumstances in which we automatically receive and collect information from you. The most common sources of this information include:

Cookies

Pixel tags or clear Graphics Interchange Format files, known as GIFs

Your Device or Apps you use

Wireless networks operated by Barnes & Noble in retail stores

Business partners, contractors, shared databases, and other third parties who may occasionally share information with us, including UltraViolet™, etc.

For your convenience, we have provided a summary description of each of these circumstances below.

Information automatically collected when you visit Barnes & Noble Websites

When you visit a Barnes & Noble Website, we automatically record information that your browser sends us. For example, we may receive and collect: the name of the domain and host from which you access the Internet; the Internet Protocol (IP) address of the computer you are using; the date and time you access the Barnes & Noble Website; and the Internet address of the website from which you linked directly to the Barnes & Noble Website. We may also collect information regarding search queries run on the Barnes & Noble Website. We use this information to monitor the usage of the Barnes & Noble Websites and as necessary for our business.

Information collected using cookies

Like many online retailers, we and/or our third party providers use cookies to recognize you as you use or return to the Barnes & Noble Websites. This is done so that we can personalize and enhance your browsing and shopping experience. "Cookies" are small files that we place on your computer's hard drive to collect information about your activities on a Barnes & Noble Website. Cookies help us to: (1) speed navigation, keep track of items in your shopping bag and provide you with content that is tailored to you; (2) remember information you gave us so that you do not have to re-enter it; (3) determine the effectiveness of some of our and our third party partners' marketing efforts and communications; and (4) monitor the total number of visitors, pages viewed, and the total number of banners displayed.

Browsers are typically set to create cookies automatically. You can choose to have your browser notify you when cookies are being written to your computer or accessed, or you can disable cookies entirely. If you disable cookies, however, you will not be able to place items in a Barnes & Noble Website shopping bag, and therefore you will not be able to place an order with us online. Also, by not using cookies, some Barnes & Noble Website features and services may not function properly.

Additionally, we work with a third party service provider to help us better understand how you use the Barnes & Noble Websites. This third party service provider will place cookies on your computer to collect information such as how you were referred to the Barnes & Noble Website, how you navigated around the Barnes & Noble Website, what you purchased and what traffic is driven by banner ads and emails. This information will help us to better serve you and provide you with more personalized information and product offerings. We do not allow this third party service provider to collect your credit card information, e-mail address or password information. This third party service provider may only use your personal information to perform services for us and may not share your personal information with anyone else, or use it for any other purpose, except on an aggregated, non-personally identifiable basis. For more information, please view this third party's privacy policy.

You may choose to opt-out of this third party's analysis of your browsing and purchasing behavior on such Barnes & Noble Website. To do so, please [click here](#).

By doing business with or interacting with Barnes & Noble, you consent to

the use of the tracking technologies as described above.

Information collected in connection with marketing e-mails and using pixel tags or clear GIFs

To help us understand the effectiveness of certain of our communications and marketing efforts, we may use sensing technologies that use pixel tags or clear GIFs (which are also called web beacons). These technologies allow us to determine the effectiveness of our e-mail and advertising and marketing efforts. For this purpose, we tie the pixel tags and clear gifs to personally identifiable information. We may also collect information regarding the links within such marketing materials that you click on and purchase statistics regarding items you buy following receipt of such marketing.

Information collected from Devices or Apps

When you use a Device or our Apps, we automatically collect information through the Device or App when it is connected to the Internet. For example, we may receive and collect information concerning device registration, settings, usage, firmware version, signal strength, search queries, network interaction, the name of the network from which you access the Internet, the Internet Protocol (IP) address of the device you are using, unique device identifiers (UDIDs), downloads, sideloaded content, configuration, or service information relating to any malfunction of the Device or App. This may include information regarding your reading or, in the case of videos, viewing behavior on such Devices or Apps, such as books or videos opened, page turns, bookmarks, annotations, or customer reviews.

In addition, to provide location-based services on Devices or through Apps, Barnes & Noble and third party application providers may automatically collect real-time geographic location information or other location-based information about you and your Device or other mobile device on which the App is installed. This information may be used to provide certain device or application functionality or to offer, provide and improve products and services.

For more information about information collected through Devices and Apps, please read the Barnes & Noble Mobile Privacy Supplement.

Information collected from outside sources

We may collect personal information and other information about you from business partners, contractors and other third parties. This includes, but is not limited to, instances in which you affirmatively authorize third parties to provide us with information. Such information may be collected via, but not limited to, software applications that you download or use on your Device or that otherwise allow you to interact with the Barnes & Noble enterprise. For example, if you link your UltraViolet account with your account with us, we may collect certain information regarding movies or television shows you have purchased from third parties as part of your UltraViolet collection of digital content.

If you rent or purchase textbooks from Barnes & Noble College Booksellers, we may also collect information about you from your college or university. This information may be used for a number of purposes including providing our services to you, sending postal mail and e-mail, fulfilling orders, removing repetitive or unnecessary information from customer lists, analyzing data, providing marketing support, customer authentication, providing search results and links (including paid listings and links), providing customer service and processing credit card payments.

Information collected from UltraViolet

If you link your account with us to your UltraViolet account, we may collect information regarding UltraViolet-enabled movies and television shows you have purchased from third party retailers. This will enable you to download or stream movies or television shows from us that you have purchased from us from such third parties. For more information, please view the UltraViolet privacy policy located at www.uvvu.com.

Information collected from visitors from outside the United States

If you are visiting Barnes & Noble or a Barnes & Noble Website from outside the United States and provide us with personal information, please note that your personal information will be transferred, stored and processed within the United States. Additionally, Barnes & Noble may transfer your personal information to other countries, for example because a server or third party service provider is located there. The data protection and privacy laws of other countries, including the United States, may not afford you the same level of protection as those in your own country. Barnes & Noble will take appropriate steps to protect your information.

By doing business or interacting with Barnes & Noble, you are consenting to the transfer, storage, and processing of your personal information to and within facilities located in the United States and other facility locations selected by Barnes & Noble.

4. HOW DO WE USE PERSONAL INFORMATION?

Barnes & Noble uses your personal information to provide you with a superior customer experience and, as necessary, to administer our business. For example, we use your personal information to:

Provide you with products, services and information, including product and subject recommendations and other information in Barnes & Noble retail stores, through e-mails from the Barnes & Noble enterprise, and on the Barnes & Noble Websites and online interactive communities, your Devices, Digital Content, Apps and software, and display associated content and advertising;

Administer our programs;

Provide customer service;

Conduct research and perform analysis in order to measure, maintain, protect, develop and improve our products and services;

Administer sweepstakes, contests, promotions or surveys and provide appropriate notifications;

Communicate with you about special offers, events, or new products or services that may be of interest to you;

Customize and enhance the Barnes & Noble Websites and advertising;

Make communications necessary to notify you regarding security, privacy, and administrative issues; and

Manage our business.

5. WITH WHOM DO WE SHARE PERSONAL INFORMATION?

Protecting the privacy and security of your personal information is a priority at Barnes & Noble. Barnes & Noble DOES NOT SELL OR RENT YOUR

PERSONAL INFORMATION TO THIRD PARTIES provided that your data may be transferred to or shared with a third party as part of a sale, merger, or acquisition of Barnes & Noble or one of its affiliates (see "Sales, Mergers, and Acquisitions" below).

Barnes & Noble shares your personal information with:

Other entities within the Barnes & Noble enterprise. Many of our customers purchase items from us both online and in our stores. In addition, Barnes & Noble owns and operates other operations and businesses and we share and use your personal information with our other business units and operations and may continue to do so even after such business units or operations are sold. We share your personal information within the Barnes & Noble enterprise (i) to help to ensure that you have a superior shopping or browsing experience no matter where you choose to shop or browse with us, (ii) to provide customer support and related services, (iii) to provide other services and products that might be of interest to you, (iv) to provide and/or manage certain cross-company promotions or programs (e.g. to honor your Barnes & Noble Member Program rewards), (v) to facilitate purchases of Devices or content (or other products) from within the Barnes & Noble enterprise, (vi) if we provide you with in store services and/or customer support for Devices and content, (vii) in connection with our commercial arrangements and business relationships within the Barnes & Noble enterprise, and (viii) as otherwise reasonably determined to provide a comprehensive retail and digital content experience for our joint customers.

Partners . We provide personal information to our partners that provide product and service offerings or technologies that we think may be of interest to you. Our partnerships may also result in products and services that allow you to publish Digital Content that may be shared with our partners. We may offer you the opportunity to opt-in to receive information from our partners, which may result in your personal information being shared with our partners. The use of your information by our partners and their vendors and contractors will be subject to the privacy policies of our vendors once such information is transferred.

Microsoft Corporation. As part of our ongoing partnership with Microsoft Corporation ("Microsoft"), we may share your personal information with Microsoft and its affiliates and subsidiaries under certain circumstances. If you consent to such sharing (including, without limitation, in connection with using a Microsoft account), Microsoft may use your personal

information to provide products and services to you related to your Device, Digital Content and Barnes & Noble and Microsoft services and products. These products and services may include, without limitation: (i) to enable purchases of Digital Content by you and provide Digital Content to you, (ii) to allow you to access, view and consume your Digital Content and provide related services to you, (iii) to publish certain Digital Content created by users and customers (as may be more fully set forth in any applicable terms or policies related thereto), (iv) to create and manage cross company user accounts and applications related to your Devices and Digital Content, including in connection with the integration between the Microsoft user and ID systems and Microsoft commerce platform and our products and services, (v) to offer joint product or service offerings, (vi) to provide customer support, (vii) to provide information to you regarding Microsoft product and service offerings, and (viii) as otherwise reasonably determined by us. Please note that any information provided to Microsoft will be treated by Microsoft in accordance with its applicable privacy policies in effect from time to time. Please refer to Microsoft's website for information related to Microsoft's collection, use and sharing of information that it obtains.

Service providers, subcontractors and agents who perform services on our behalf. We provide personal information to third party service providers, subcontractors, and agents that work under contract on our behalf to provide certain services. These third parties do not have the right to use personal information we provide to them in any way that is not authorized by Barnes & Noble. They are contractually obligated to use such information only as necessary to assist us and to maintain the confidentiality and security of such information.

Third party providers of products and services. These are third parties that provide products and services you may purchase or request from or through us. These third parties do not have the right to use personal information we provide to them in any way that is not authorized by Barnes & Noble. They are contractually obligated to use such information only as necessary to assist us and to maintain the confidentiality and security of such information.

UltraViolet. If you purchase UltraViolet-enabled movies or television shows from us, we may provide you with the opportunity to link your account with us to your UltraViolet account. This will enable you to download or stream movies or television shows that you have purchased from us from third party retailers who are part of the UltraViolet content ecosystem. For more information, please view the UltraViolet privacy policy located at

www.uvvu.com.

Third Party Membership Programs. We work with certain third party entities to help them administer their own membership or rewards programs by providing them with purchasing information about their customers who make purchase from Barnes & Noble. We disclose only the information that is necessary to make these third party programs work and to support your membership in them. This information usually includes your name and e-mail address as well as the dollar amount of the purchases made at Barnes & Noble. We require such entities to obtain your consent before we provide them with this information, which they usually do as part of their own membership or participation rules.

Credit Card Companies. Credit card transactions are handled by third party financial institutions and their vendors and contractors who receive credit card numbers and other personal information from us to verify the credit card numbers and process transactions. Although Barnes & Noble's treatment of this information is governed by this policy, the use of your information by the third party financial institutions and their vendors and contractors will be subject to their own privacy policies.

Law enforcement officials and as required by law. Barnes & Noble may release personal information to third parties when we determine, in our judgment, that it is necessary to (a) comply with the law, regulation, legal process, or enforceable governmental request; (b) enforce or apply the terms of any of our policies or user agreements; or (c) protect the rights, property or safety of Barnes & Noble, our employees, our customers, users, or others.

Co-branded offerings. In some instances, we may offer a service or feature that is co-branded by Barnes & Noble and a third party company. If you provide information in connection with such co-branded service or feature, that information may be shared between Barnes & Noble and the third party. Although Barnes & Noble's treatment of the information is governed by this policy, the third party's treatment of your information will be subject to the third party's privacy policy.

Other websites operated by Barnes & Noble. In some instances, we may operate a Barnes & Noble Website on behalf of a third party. If you provide information in connection with a Barnes & Noble Website operated by Barnes & Noble on behalf of a third party, that information may be shared between Barnes & Noble and the third party. Although Barnes & Noble's

treatment of the information is governed by this policy, the third party's treatment of your information will be subject to the third party's privacy policy.

Your college or university. Barnes & Noble College Booksellers may offer the ability to rent or purchase products or services through financial aid, student loans or similar programs. We may share your information with your college or university, lenders or other providers in order to process payments if you choose these payment methods.

Sales, Mergers, and Acquisitions. If Barnes & Noble becomes involved in a merger, acquisition, restructuring, reorganization, or any form of sale or other disposition of some or all of its assets, personal information and your transaction history may be provided to the entities and advisors involved subject to a confidentiality agreement, and we will provide notice to you at your email address on file before any personal information is finally transferred and becomes subject to a different privacy policy. Such a transaction may involve us (i) retaining the right to continue to use transferred personal information in addition to the right of the transferee to use such information, and (ii) engaging in additional transfers of personal information (including new personal information) with the transferee from time to time following such a transaction.

To Fulfill Your Requests For Products or Services

In connection with some Digital Content orders, in order to complete your transaction, we may need to forward your name, magazine, catalog or newspaper order, email address, and shipping or billing address to our content providers. The content providers will share your magazine, catalog or newspaper order information as well as your name and address, or name and e-mail and shipping or billing addresses with the magazine, catalog or newspaper publisher and magazine, catalog or newspaper circulation auditors. Magazine, catalog and newspaper publishers may use this information to fulfill your order and for other purposes. Your credit card information will not be shared with them.

In connection with purchases of certain Digital Content, we may need to forward information about you (including, for example, your Internet Protocol (IP) address) to the Digital Content provider in order to enable you to download or purchase Digital Content from or through that provider. In

each such instance, the Digital Content provider is obligated to use such information in accordance with its own privacy policy and applicable law.

6. HOW DO WE SECURE PERSONAL INFORMATION?

We take significant and appropriate security measures, including physical, technological and procedural measures, to help to safeguard your personal information and to prevent unauthorized access and disclosure. In addition, we use industry-standard technology, such as Secure Sockets Layer (SSL) encryption technology in the transmission of certain sensitive personal information, designed to prevent unauthorized persons from gaining access to your personal information, and, as technology develops, we intend to take additional measures to improve security.

We want you to feel confident whenever you visit us in our stores, on the Barnes & Noble Websites, or use Barnes & Noble Devices or Apps. While we are focused on the security of your personal information and follow strict standards, processes and procedures that are designed to protect your personal information, you must remember that the Internet is a global communications vehicle open to threats, viruses and intrusions from others and so we cannot promise, and you should not expect, that we will be able to protect your personal information at all times and in all circumstances.

7. HOW DO WE RESPECT YOUR CHOICES ABOUT YOUR PERSONAL INFORMATION?

When you interact with Barnes & Noble in certain ways, you may be eligible to receive certain marketing-related and promotional communications as well as special offers (collectively "Promotional Communications") from Barnes & Noble that may include advertisements from third parties. The most common of these circumstances include:

- Making a purchase or download or placing an order
- Creating an account
- Joining a Barnes & Noble online community
- Enrolling in our Member program
- Signing up for Newsletters
- Signing up with eBookstores operated by Barnes & Noble
- Entering a sweepstakes or a contest

Where you have consented to receiving Promotional Communications from the Barnes & Noble enterprise, you may choose to opt-out at any time by

following the instructions below.

For Barnes & Noble Booksellers: At any time, a customer may choose to opt-out of the receipt of any promotional communications by clicking on the opt-out link provided at the bottom of each e-mail and following the instructions, or by e-mailing our Customer Service Department by clicking [here](#).

For Barnes & Noble College Booksellers: At any time, a customer may choose to opt-out of the receipt of any promotional communications by clicking on the opt-out link provided at the bottom of each e-mail and following the instructions.

For barnesandnoble.com: Any account holder can opt-out of the receipt of any promotional communications by logging onto their Account and following the instructions under "Communication Preferences". If you do not have a barnesandnoble.com account, you may opt-out of the receipt of such communications by selecting the link at the bottom of any promotional e-mail communication and following the instructions.

For Barnes & Noble Member program: Any Member can opt-out of the receipt of any Member program promotional communications by logging on to their barnesandnoble.com Account and following the instructions under "Change Your Communication Preferences." If you do not have a barnesandnoble.com account, you may opt-out of the receipt of such communications by selecting the link at the bottom of any promotional e-mail communication and following the instructions.

For Barnes & Noble Educator program: Any Educator can opt-out of the receipt of any promotional communications by logging on to their barnesandnoble.com Account, clicking "Communication Preferences," and following instructions under "Change Communication Preferences." If you do not have a barnesandnoble.com account, you may opt-out of the receipt of such communications by clicking on the opt-out link provided at the bottom of each e-mail and following the instructions. If no opt-out link exists reply to the email with the word "Remove" in the subject line, or by e-mailing our Customer Service Department by clicking [here](#).

For the Barnes & Noble Gift Card website: Any customer of the Barnes & Noble Gift Card website who receives promotional communications can

opt-out of the receipt of any such communications by clicking on the opt-out link provided at the bottom of each e-mail and following the instructions or by sending an e-mail to giftcardprefs@bn.com.

For SparkNotes.com: Any SparkNotes.com customer who does not wish to receive promotional product and service communications can choose to remove their contact information from SparkNotes.com contact list by following the instructions listed on www.sparknotes.com.

For Sterling Publishing: If, at any time, you wish to no longer receive ENewsletters from Sterling Publishing Co., Inc., Pixiq, Lark Crafts, and/or FlashKids.com, please click on the link for opt-out instructions contained at the bottom of each ENewsletter. You can change your personal information and your communication preferences when you follow this link.

Even if you opt-out of receiving Promotional Communications, you may continue to receive e-mails relating to order confirmations, back order notifications, membership information, and/or other business-related communications.

Note for Community Users

Barnes & Noble provides you with the ability to access, correct, change or request deletion of the personal information in your community profile(s) at any time by following the instructions below. We will respond to all access requests within 30 days. However, you cannot currently change your profile name although you may re-register and choose a new name. To review or modify your profile information, please click [here](#). Please be advised that Barnes & Noble may archive information it collects on its community and visitors. Additionally, by participating in the Barnes & Noble online interactive community, and using Barnes & Noble's various interactive offerings, including, the submission of customer reviews, or participation in the Barnes & Noble Book Clubs, you agree to receive communications from Barnes & Noble, other users, and moderators related to the provision of these services.

You may access, correct or change the personal information in your community profile(s) on Lark Crafts and Pixiq at any time. To review or modify your profile information, log in to larkcrafts.com or pixiq.com. Additionally, by using Lark Crafts and/or Pixiq's various interactive

offerings, you agree to receive communications from Barnes & Noble, Lark Crafts, Pixiq, other users, and moderators related to the provision of these services.

You may also access, correct or change the personal information in your community profile(s) on SparkNotes.com at any time, except to change your username. To review or modify your profile information, log in to SparkNotes.com. Additionally, by using SparkNotes.com's various interactive offerings, you agree to receive communications from Barnes & Noble, SparkNotes, other users, and moderators related to the provision of these services.

We will retain your information for as long as your account is active or as needed to provide you with services. We will retain and use your information as necessary to comply with our legal obligations, resolve disputes, and enforce our agreements.

8. WHOM DO I CONTACT IF I HAVE QUESTIONS OR CONCERNS?

If you have any questions about this Privacy Policy, please contact us at privacy@barnesandnoble.com or call customer service at 1-800-THE-BOOK [1-800-843-2665].

9. WHAT OTHER INFORMATION ABOUT BARNES & NOBLE'S PRIVACY POLICY WOULD I WANT TO KNOW?

Barnes & Noble Policies for Minors, including children under the age of 13.

Barnes & Noble is committed to complying with all applicable laws and regulations regarding the collection, storage and use of personal information of Minors under the age of 13, including the Children's Online Privacy Protection Act of 1998.

Except as noted below : (i) our products and services are marketed for and directed towards purchase by adults or with the consent of adults; (ii) Individuals under the age of 18 ("Minors") are not permitted to use Barnes & Noble Websites or Apps without the supervision of a parent or legal guardian; (iii) we do not knowingly collect or solicit personal information from children under the age of 13 or knowingly allow such persons to register for an online account or to post personal information on the Barnes & Noble Websites; and (iv) should we learn that someone under the age of 13 has provided any personal information to or on any of the Barnes &

Noble Websites, we will remove that information as soon as possible.

If we direct certain Devices features or Apps to Minors who may be under the age of 13, we require the parent or legal guardian to provide verifiable consent to such Minor's use of our products and services and our collection of personal information in connection with such use. In these cases, we require validation of credit card information already on file with us and we will then e-mail the master account holder to give notice that we have received such consent. If no credit card information is on file, or if the credit card is invalid, we will not register that you have consented. With your consent, we may collect, use and share information regarding Minors under the age of 13 consistent with this Privacy Policy. Please see the sections above for detailed information on how we do so.

If you are a parent or guardian who believes that your Minor under the age of 13 has submitted personal information or other information to us without your consent, please contact privacy@barnesandnoble.com. Once we verify that you are the parent or legal guardian, at your request, we will promptly provide to you information regarding what, if any, personal information we have collected about your child and how it has been used or shared. We will, at your request, remove personal information about your child from its database and instruct our affiliates and third party partners to do the same.

Your access rights and updates to your information

You have the right to ask us not to use your personal information for direct marketing purposes. You also have the right to request a copy of the personal information that we hold about you and to have any inaccuracies rectified. Please note that, as permitted by law, we may charge a nominal fee for information requests and may require you to prove your identity. Following a request, we will use reasonable efforts to supply, correct or delete personal information about you in our files. Please address written requests and questions about this to privacy@barnesandnoble.com. We will respond to all access requests within 30 days.

Third Party Applications: We are not responsible for third party applications you may download and install on your Devices. While Barnes & Noble uses reasonable efforts to ensure that applications offered through our Application Stores meet compatibility and safety testing criteria, this Privacy Policy does not cover any data collection, use, or sharing by mobile device manufacturers, network service providers or third party application

providers, unless such party has entered into a separate agreement with us as a business partner to share your personal information. In such a case, we would require our business partner to disclose their data collection and sharing practices with you, and obtain your prior affirmative consent. Our use of any shared data under such an agreement would then be governed by our privacy policy. All other collection of personal information by a third party application provider shall be subject to the terms of such provider's privacy policy.

Third Party Websites and Web-Tracking

We may partner with third party distributors of advertisements to deliver advertisements about our products and services when you visit other websites across the Internet. These distributors of advertisements may place or access cookies on your computer which help us provide you with advertisements that are customized to your particular product preferences and improve the effectiveness of our marketing efforts. If you would like to learn more about the third party advertisers that may be aware of the fact that you visit Barnes & Noble Websites, and to understand your choices about having such advertisers' cookies turned off, please visit www.networkadvertising.org. Please note this does not opt you out of being served advertising. You will continue to receive generic ads.

Please understand that Barnes & Noble does not control these third party cookies, and you should check the privacy policy of the Internet advertising company or advertiser to see whether and how it uses cookies.

Whenever you click on links and banners on any Barnes & Noble Website that take you to a third party website, you will be subject to the third party's privacy policies, not Barnes & Noble's. This Privacy Policy applies solely to information collected by the Barnes & Noble enterprise.

In addition, our website includes links to other websites whose privacy practices may differ from those of Barnes & Noble. If you submit personal information to any of those sites, your information is governed by their privacy policies. We encourage you to carefully read the privacy policy of any website you visit.

Safe Harbor

Barnes & Noble's privacy policy has been reviewed by TRUSTe for

transparency, accountability and choice for our collection and use of your personal information. If you have questions or complaints regarding our privacy policy or practices, please contact us at privacy@barnesandnoble.com. If you are not satisfied with our response you can contact TRUSTe here.

Barnes & Noble complies with the U.S. - E.U. Safe Harbor framework and the U.S. - Swiss Safe Harbor framework as set forth by the U.S. Department of Commerce regarding the collection, use, and retention of personal data from, respectively, European Union member countries and Switzerland. Barnesandnoble.com LLC has certified that it adheres to the Safe Harbor Privacy Principles of notice, choice, onward transfer, security, data integrity, access, and enforcement. To learn more about the Safe Harbor program, and to view Barnes & Noble's certification, please visit <http://www.export.gov/safeharbor/>

Terms of Use

Any dispute between you and us regarding the privacy of your personal information is subject to this Privacy Policy and the terms of use or terms of service applicable to our product or service you use, including limitation on damages, resolution of disputes, and application of the law of the State of New York.

Contact Information

The address for Barnes & Noble is 122 Fifth Avenue, New York, NY 10011. The address for NOOK Media is 76 Ninth Avenue, New York, NY 10011

III. LOWES.COM

Lowe's respects your concerns about privacy. This Privacy Statement applies to personal information we collect on Lowe.com and other official Lowe's sites (including mobile sites) that link to this Privacy Statement, the Lowe's Career Website hosted on brassring.com, official Lowe's survey forms and sweepstakes entry forms on third-party sites and applications that link to this Privacy Statement, official Lowe's mobile applications that link to this Privacy Statement (the "Mobile Applications"), and from use of the in-store Wi-Fi network (the "Store Wi-Fi") (collectively, the "Site"). The credit-related areas of the Site are covered by different privacy notices that appear on the credit application pages or that you receive in connection with your credit accounts, as applicable; the term "Site" does not refer to these areas.

The term "Site" also does not refer to the Weekly Ads pages at lowes.shoplocal.com, the Iris Smart Home website www.irissmarthome.com or the login page for that website on Lowe.com. In some cases, a website or online service covered by this Privacy Statement may provide additional detail about privacy practices specific to the website or online service.

This Privacy Statement describes the types of personal information we collect on the Site, how we may use that information and with whom we may share it. The Privacy Statement also describes the measures we take to protect the security of the personal information. We also tell you how you can reach us to ask us to update your preferences regarding how we communicate with you or answer any questions you may have about our privacy practices.

INFORMATION WE COLLECT

You may choose to provide us with personal information (such as name, contact details and payment information), such as:

Contact information, such as your name, address, telephone number, and email address, and your title or occupation.

Login and access credentials (such as username and password) for Lowe's accounts.

Payment information, such as your payment card number and expiration date.

Date of birth.

The geolocation of your device (such as if you opt to use the "Find Near Me" feature of the mobile-optimized portion of our websites or our Mobile Applications).

The unique ID number associated with certain Lowe's accounts.

Certain information about your Lowe's purchases returns or exchanges.

Personal information in communications and other content you submit, such as photographs, product information and details about your projects and property.

Social media IDs, such as for Facebook or Twitter

Personal information you submit in your capacity as an independent contractor of Lowe's, such as gender, ethnicity and identification numbers such as your NRDS ID number.

Personal information you submit in connection with a job application on the Career Website, such as your name, contact information, Social Security Number, date of birth, employment status, employment history, education information, references, rÃ©sumÃ©, immigration status and ability to work

legally in the United States, driver license information, personal or family employment affiliation with Lowe's, criminal record, and, on a voluntary basis for our Equal Employment Opportunity compliance purposes for jobs located in the US, gender, race and ethnic background.

[Back to Top](#)

INFORMATION WE COLLECT BY AUTOMATED MEANS

When you use our Site, we may collect certain information by automated means, using technologies such as cookies, Web server logs, Web beacons and JavaScript.

Cookies are files that websites send to your computer or other Internet-connected device to uniquely identify your browser or to store information or settings on your device. Our Site may use HTTP cookies, HTML5 cookies, Flash cookies and other types of local storage (such as browser-based or plugin-based local storage). Your browser may tell you how to be notified when you receive certain types of cookies and how to restrict or disable certain cookies. You also may be able to delete your Flash cookies or adjust your Flash cookie settings by visiting the Adobe Flash Website Storage Settings Panel and Global Storage Settings Panel. Please note, however, that without cookies you may not be able to use all of the features of our Site or other websites and online services.

In conjunction with the gathering of information through cookies, our Web servers may log information such as your device type, operating system type, browser type, domain, and other system settings, as well as the language your system uses and the country and time zone where your device is located. The Web server logs also may record information such as the address of the Web page that referred you to our Site and the IP address of the device you use to connect to the Internet. They also may log information about your interaction with this Site, such as which pages you visit. To control which Web servers collect information by automated means, we may place tags on our Web pages called "Web beacons," which are small files that link Web pages to particular Web servers and their cookies. We also may send instructions to your device using JavaScript or other computer languages to gather the sorts of information described above and other details your interactions with the Site.

We may use third-party Web analytics services on our Site, such as those of Google Analytics. These service providers use the technology described above to help us analyze how users use the Site. The information collected by

the technology (including your IP address) will be disclosed to these service providers, who use the information to evaluate your use of the Site. To learn about opting out of Google Analytics, please [click here](#).

We may use the information collected through automated means to provide a better tailored shopping experience and for market research, data analytics and system administration purposes, such as to determine whether you've visited us before or are new to the Site, and for compliance with our legal obligations, policies and procedures. We also may use this information to target custom content and ads to you on this and other websites, including as described in the Interest-Based Advertising section below.

[Back to Top](#)

INFORMATION COLLECTED AUTOMATICALLY BY THE MOBILE APPLICATIONS

If you elect to install the Mobile Applications, the information we collect may include the following:

Your geographic location, including within a Lowe's store and the surrounding area.

Information about your use of the Mobile Applications.

The type of device you use and its operating system.

Identification details of your device (e.g., unique device identifier).

IP address.

Your use of mobile coupons in a Lowe's transaction.

This information will allow push notifications and other targeted marketing designed specifically for your shopping preferences such as special offers based upon areas in which you may be shopping in the store, and shopping lists with specific items located for your convenience when you are shopping a particular store, as well as for other in-store mapping and routing services. It also may be used for the other purposes specified in this Privacy Statement.

[Back to Top](#)

INFORMATION COLLECTED AUTOMATICALLY BY STORE WI-FI

If you elect to use the free Store Wi-Fi available at Lowe's stores, we will collect information such as the following:

The URLs and content of the pages you visit on any website (e.g., on

yahoo.com or Lowes.com) using your mobile device, and, on some websites, information you submit through online forms.

The apps on your mobile device that use the Store Wi-Fi.

Your geographic locations within a Lowe's store and the surrounding area within the range of the Store Wi-Fi.

How long you use the Store Wi-Fi at particular locations.

The type of device you use and its operating system.

Identification details of your device (e.g., unique device identifier and MAC address).

Browser information and IP address.

The address of the store where you use the Store Wi-Fi.

This information will allow targeted marketing designed specifically for your shopping preferences such as specific coupons based upon the sites and pages you visited, special offers based upon areas in which you may be shopping in the store, including competitive offers based upon other websites that you may be viewing, shopping lists with specific items located for your convenience when you are shopping a particular store, as well as for other in-store mapping and routing services. It also may be used for the other purposes specified in this Privacy Statement.

[Back to Top](#)

INTEREST-BASED ADVERTISING

Data about your activities online is being collected on our Site for use in providing advertising tailored to your individual interests. You may choose whether or not to have your information collected for that purpose. This section of the Privacy Statement provides details and explains how to exercise that choice.

You may see certain ads on other websites because we participate in advertising networks administered by third parties. These networks track your online activities over time by collecting information through automated means, including through the use of the technologies described in the Information We Collect By Automated Means section above, and they use this information to show you advertisements that are tailored to your individual interests. The information they collect includes information about your visits to our Site, such as the pages you have viewed. This collection and ad targeting takes place both on our Site and on third-party websites that participate in the ad network, such as sites that feature advertisements delivered by the ad network. This process also helps us track the effectiveness of our marketing efforts.

To learn more about ad networks, including how to opt out, click [here](#).

[Back to Top](#)

HOW WE USE THE INFORMATION WE COLLECT

We may use the information we collect through the Site to:

Provide, administer and communicate with you about products, services, events and promotions (including by sending you newsletters, coupons and other marketing communications).

Process, record and track your purchases, payments and rebates.

Process, evaluate and respond to your requests, inquiries and applications.

Manage our customer information databases.

Operate registrations, including on MyLowe's and Lowes.com.

Administer contests, sweepstakes and surveys.

Create, administer and communicate with you about your accounts.

Customize your experience with our Site.

Provide you with in-store navigation and mapping services.

Allow you to find locations near you.

With respect to personal information we collect through our Career Website, evaluate your application for employment and contact you regarding possible employment.

Operate, evaluate and improve our business (including developing new products and services; managing our communications; performing market research, data analytics and data appends; determining and managing the effectiveness of our advertising and marketing; analyzing our products, services and Site; administering our Site; and performing accounting, auditing, billing, reconciliation and collection activities).

Protect against and prevent fraud, unauthorized transactions, claims and other liabilities, and manage risk exposure and quality.

Comply with and enforce applicable legal requirements, industry standards and our policies and terms, such as our Terms and Conditions of Use.

We also may use the information in other ways for which we provide notice at the time of the collection.

[Back to Top](#)

SHARING OF INFORMATION

We may share personal information we collect on the Site with certain

service providers, some of whom may use the information for their own purposes. For example, (i) our websites may feature live chat functionality, and the information gathered in the chat feature may be collected by or shared with a provider such as LivePerson, whose privacy policies are available at www.liveperson.com/policies/privacy, (ii) information you submit on our Site in connection with a product review may be collected by or shared with a product review company such as Bazaarvoice, whose privacy policy is at www.bazaarvoice.com/privacy-policy and who may publish the information in locations not affiliated with Lowe's, such as the website of the manufacturer of the product you review, and (iii) information submitted on our Careers Website is collected by or shared with the Careers Website provider, Kenexa, whose Privacy Policy is at www.kenexa.com/privacypolicy. We also may share your information among our affiliates, joint marketing partners and companies that fulfill orders placed with us, who may send you marketing information. In addition, we may share personal information we collect on the Site at your request.

We may disclose information about you (i) if we are required to do so by law, regulation or legal process, such as a court order or subpoena; (ii) in response to requests by government agencies, such as law enforcement authorities; or (iii) when we believe disclosure is necessary or appropriate to prevent physical, financial or other harm, injury or loss; (iv) in connection with an investigation of suspected or actual unlawful activity; or (v) to assist in collecting debt owed by you.

We reserve the right to transfer personal information we have about you in the event we sell or transfer all or a portion of our business or assets (including, without limitation, in the event of a reorganization, dissolution or liquidation). Should such a sale or transfer occur, we will use reasonable efforts to direct the transferee to use personal information you have provided to us in a manner that is consistent with our Privacy Statement.

[Back to Top](#)

YOUR CHOICES

Where applicable, you may amend your preferences regarding how we communicate with you by updating your profile within your Lowe's authenticated account, using the settings pages on certain websites or contacting us as described in the How to Contact Us section below.

To be removed from all of Lowe's official email, telephone and postal mail

marketing, choose one of the following options: email customercare@lowes.com and type "REMOVE FROM ALL MARKETING" in the subject line; call 1-800-44-LOWES; or send your request by mail to:

Lowe's Customer Care
P.O. Box 1111
North Wilkesboro, NC 28659

For any of these options, please include your name, address, phone number and email address in the request, and let us know how you provided us with the information.

[Back to Top](#)

ACCOUNT PROFILE AND COMMUNICATIONS PREFERENCES GENERALLY

You may access the profile page of certain accounts you maintain on our Site to modify certain personal information associated with your profile and indicate certain communications preferences.

[Back to Top](#)

EMAIL

Where applicable, you may amend your preferences regarding how we communicate with you through email by clicking on the "unsubscribe" link in an email you receive from us, by updating certain subscriptions within certain Lowe's authenticated accounts, within settings pages or by contacting us as described in the How to Contact Us section below. To stop receiving newsletters from Lowe.com, you also may email customercare@lowes.com with the phrase "REMOVE FROM NEWSLETTER LIST" in the subject line, and we will apply your preference to the email address from which you send us the request.

[Back to Top](#)

INTEREST-BASED ADVERTISING

To exercise your preferences regarding the collection of data for targeted advertising, please follow the directions in the Interest-Based Advertising section above.

[Back to Top](#)

STORE WI-FI

To stop the collection of information via Store Wi-Fi, disconnect from the Store Wi-Fi network. Please note that disconnecting from the Store Wi-Fi network will not prevent you from taking advantage of our Mobile Applications' collection of your location. To deactivate that collection of information, follow the instructions in the Mobile Applications section below.

[Back to Top](#)

MOBILE APPLICATIONS

To stop collection of information by the Mobile Applications, delete the applications from your device. To stop only the collection of information from the geolocation services available on your device (such as your device's GPS functionality), use your device's settings to deactivate the Mobile Applications' access to those services. Please note that deactivating this access will not prevent you from taking advantage of our Store Wi-Fi's collection of your location. To deactivate that collection of information, follow the instructions in the Store Wi-Fi section above.

[Back to Top](#)

THE LOWE'S PRIVACY AND SECURITY STATEMENT

To exercise a choice under a prior version of this statement, contact us as described in the [How to Contact Us](#) section below.

[Back to Top](#)

NOTICE TO CALIFORNIA CUSTOMERS

Subject to certain limitations under California Civil Code Â§ 1798.83, if you are a California resident, you may ask us to provide you with (i) a list of certain categories of personal information that we have disclosed to certain third parties for their direct marketing purposes during the immediately preceding calendar year and (ii) the identity of certain third parties that received personal information from us for their direct marketing purposes

during that calendar year. To make such a request, please contact us as follows:

Lowe's Customer Care

Attn: Privacy Team "California Marketing Choices" CON8

P.O. Box 1111

North Wilkesboro, NC 28656

1-800-445-6937

[Back to Top](#)

HOW WE PROTECT PERSONAL INFORMATION

We maintain administrative, technical and physical safeguards to protect the personal information you provide on our Site against accidental, unlawful or unauthorized destruction, loss, alteration, access, disclosure or use and other unlawful forms of processing.

[Back to Top](#)

LINKS TO OTHER WEBSITES

This Site contains links to other websites for your convenience and information. These websites may be operated by companies not affiliated with us. Linked websites, including those operated by Lowe's, may have their own privacy notices, which we strongly suggest you review if you visit them. We are not responsible for the content of any websites that we do not control, any use of those websites or the privacy practices of those websites.

[Back to Top](#)

UPDATES TO OUR PRIVACY STATEMENT

This Privacy Statement may be updated periodically and without prior notice to you to reflect changes in our personal information practices or relevant laws. We will post a notice on Lowes.com and certain other parts of the Site to notify you of any significant changes to our Privacy Statement and indicate at the top of the Privacy Statement when it was updated.

[Back to Top](#)

BBB ONLINE

We participate in the Council of Better Business Bureau's BBBOOnLine program. Further information about this program and its dispute resolution process is available at the Better Business Bureau's website.

[Back to Top](#)

HOW TO CONTACT US

If you have any questions or comments about this Privacy Statement, or if you would like us to update information we have about you or your preferences, please contact us as indicated below. When you contact us about personal information we hold about you, please let us know where you provided us with the information.

Lowe's Customer Care
Attn: Privacy Team " CON8

P.O. Box 1111

North Wilkesboro, NC 28656

1-800-445-6937

customercare@lowes.com

IV. OVERSTOCK.COM

Overstock.com Privacy and Security Policy

Overstock.com, Inc. ("Overstock.com," "Overstock," "we," "us," or "our" as appropriate) has created the following Privacy and Security Policy ("Privacy Policy") to inform you of the following: what information we collect, when we collect it, how and why we collect it, how we use it, how we protect it, how we share it and what choices you have regarding our collection and use of this information.

This Privacy Policy applies to the following sites:

<http://www.overstock.com>, <http://www.o.info>, and <http://www.o.biz> (collectively and individually, the "Sites"). Entering the Sites constitutes your acceptance and agreement to the terms contained in this Privacy Policy. If you do not agree with the terms set forth in this Privacy Policy, please do not enter the Sites.

WHAT INFORMATION WE COLLECT

Types of Information We Collect

- Your name
- Your billing and delivery address
- Your e-mail address
- Your phone (or mobile) number
- Your credit/debit card number
- Information on how you are using the Sites
- Your purchase/return/exchange information

We Collect Information When

- You purchase, order, return, exchange or request information about our products and services from the Sites or mobile applications.
- You create an Overstock.com account
- You connect with Overstock.com regarding customer service via our customer service center, or on social media platforms.
- You visit the Sites or participate in interactive features of the Sites or mobile applications.
- You use a social media service, for example, Overstock.com's Facebook page or YouTube channel.
- You sign up for e-mails, mobile messages, or social media notifications from Overstock.com.
- You enter a contest or sweepstakes, respond to one of our surveys, or participate in a focus group.
- You provide us with comments, suggestions, or other input
- You interact with any of the Sites through your computer, tablet or mobile device

HOW AND WHY WE COLLECT THE INFORMATION

Technologies Used

We may use tracking pixels/web beacons, cookies and or other technologies

to receive and store certain types of information. This information includes Internet Protocol (IP) addresses, browser information, Internet Service Provider (ISP), operating system, date/time stamp and clickstream data. This information helps us customize your website experience and make our marketing messages more relevant. It also allows us to provide features such as storage of items in your cart between visits. This includes Overstock.com content presented on other websites or mobile applications. In order to provide the best customer experience possible, we also use this information for reporting and analysis purposes, such as how you are shopping our website, performance of our marketing efforts, and your response to those marketing efforts.

Third Party Cookies

We allow third-party companies to collect anonymous information when you visit the Sites and to use that information to serve ads for Overstock.com products or services or for products or services of other companies when you visit the Sites or other websites. These companies may use anonymous information (e.g., navigational, non-personally identifiable information, click stream information, browser type, time and date, subject of advertisements clicked or scrolled over, etc.) during your visits to the Sites and other websites in order to provide advertisements about our goods and services likely to be of interest to you. These parties may use a cookie or a third party web beacon, or other technologies, to collect this information.

To opt out of third-party vendors' cookies, see the "What Choices Do You Have?" section of this privacy policy for instruction how to do so.

User Experience Information

In order to improve customer online shopping experience, help with fraud identification, and to assist our customer care representatives in resolving issues customers may experience in completing online purchases, we use tools to monitor certain user experience information, including, but not limited to, login information, IP address, data regarding pages visited, specific actions taken on pages visited (e.g., information entered during checkout process), and browser information.

Information from Other Sources

We may also obtain information from companies that can enhance our existing customer information to improve the accuracy and add to the information we have about our customers (e.g., adding address information). This may improve our ability to contact you and improve the relevancy of our marketing by providing better product recommendations or special offers that we think will interest you.

Public Forums

Any information you submit in a public forum (e.g., a blog, chat room, or

social network) can be read, collected, or used by us and other participants, and could be used to personalize your experience. You are responsible for the information you choose to submit in these instances.

Mobile Privacy

Overstock.com offers mobile applications (commonly known as "apps") that allow you to shop online, check product availability, learn about sales events, or receive other information from Overstock. All information collected by Overstock.com via our mobile application is protected by this Privacy Policy. Although you do not have to provide your location information to Overstock.com to use our mobile applications, our services require a zip code to function. If you have questions about location and notification privacy, please contact your mobile service provider or the manufacturer of your device to learn how to adjust your settings.

HOW WE USE THE INFORMATION WE COLLECT

Product and Service Fulfillment

Fulfill and manage purchases, orders, payments, returns/exchanges, or requests for information, or to otherwise serve you

Provide any requested services

Administer sweepstakes and contests

Marketing Purposes

Deliver coupons, mobile coupons, newsletters, receipt messages, e-mails, and mobile messages

Send marketing communications and other information regarding products, services and promotions

Administer promotions

Internal Operations

Improve the effectiveness of the Sites, mobile experience, and marketing efforts

Conduct research and analysis, including focus groups and surveys

Perform other business activities as needed, or as described elsewhere in this policy

Fraud Prevention

To prevent fraudulent transactions, monitor against theft and otherwise protect our customers and our business

Legal Compliance

To assist law enforcement and respond to subpoenas

HOW WE PROTECT THE INFORMATION WE COLLECT

Security Methods

We maintain technical, administrative, physical, electronic and procedural safeguards to protect the confidentiality and security of information transmitted to us. To guard this information, the Sites use Secure Sockets Layer (SSL). SSL encrypts your credit card number, name and address so only Overstock.com is able to decode the information.

E-mail Security

Please note that e-mail is not encrypted and is not considered to be a secure means of transmitting credit card information. "Phishing" is a scam designed to steal your information. If you receive an e-mail that looks like it is from us asking you for certain information, do not respond. Though we might ask you your name, we will never request your password, credit card information or other information through e-mail.

Information About Children

We recognize the particular importance of protecting privacy where children are involved. We are committed to protecting children's privacy and we comply fully with the Children's Online Privacy Protection Act (COPPA). Overstock.com will never knowingly request or collect personal information from any person under 13 years of age without prior verifiable parental consent. If we become aware that an individual is under the age of 13 and has submitted any information to Overstock.com, for any purpose without prior verifiable parental consent, we will delete his or her information from our files.

Additional Security

We also ask customers to carefully review their accounts and immediately report any unexpected activity to Overstock.com and their issuing bank or credit card company. We are asking all our customers to take measures to help protect information in their online accounts, including the following: Install the latest security updates and anti-virus software on your computer to help prevent malware and viruses

Reset your e-mail account passwords frequently

Use complex passwords (a minimum of 7 alpha/numeric cAsE sEnsitive characters)

Do not use the same password on more than one website

Do not share your password with others

Sign out/log off website sessions so that your session is closed and cannot be accessed by another user on the same computer, especially when using a public computer or terminal

HOW WE SHARE THE INFORMATION WE COLLECT

General Policy

We do not sell or rent customer information to third parties; except, under

limited circumstances outlined below, we may share information with third parties.

The Overstock Family

We share information we collect within the Sites and Overstock.com family, which includes all Overstock.com subsidiaries and affiliates. The Overstock.com family may use this information to offer you products and services that may be of interest to you.

Service Providers

We may share information with companies that provide support services to us (such as a printer, e-mail, mobile marketing, or data enhancement provider) or that help us market our products and services. These companies may need information about you in order to perform their functions. These companies are not authorized to use the information we share with them for any other purpose.

Legal Requirements

We may disclose information you provide to us when we believe disclosure is appropriate to comply with the law; to enforce or apply applicable terms and conditions and other agreements; or to protect the rights, property or safety of our company, our customers or others.

When You Direct Us

At your direction or request, we may share your information.

Business Transfers

If some or all of our business assets are sold or transferred, we generally would transfer the corresponding information regarding our customers. We also may retain a copy of that customer information.

WHAT CHOICES DO YOU HAVE

E-mail

Promotional E-mail

If you do not wish to receive promotional e-mails from us, [click here](#). You also have the ability to unsubscribe to promotional e-mails via the opt-out link included in each e-mail. It may take up to 10 business days before you stop receiving promotional e-mails.

Important Notices and Transactional E-mail

From time to time, we may send non-commercial electronic email messages with important information about us or the Sites to your email address. We regularly send email order confirmations and email order updates to you after you have submitted an order.

Mobile

We may distribute mobile coupons and text messages to mobile devices of customers who have requested this information via an opt-in request.

Customers will be able to opt out of a specific mobile messaging campaign.
Overstock.com Cookies

The help function of your browser should contain instructions to set your computer to accept all cookies, to notify you when a cookie is issued, or to not receive cookies at any time. If you set your computer to not receive cookies at any time, certain personalized services cannot be provided to you, and accordingly, you may not be able to take full advantage of all of the Overstock.com features.

Third Party Cookies

To opt-out of third-party vendor's cookies on other websites, visit the Network Advertising Initiative website, click here
<http://www.networkadvertising.org/choices/>.

Telephone

If you do not wish to receive promotional communication from us, call us at (800) 843-2446 to opt out. This opt out does not apply to operational communication, for example, confirmation of delivery address.

HOW DO YOU ACCESS AND UPDATE THE INFORMATION

In order to keep your information accurate and complete, you can access or update some of your information in the following ways:

If you have created an Overstock.com account, you can log in and update your account information, including contact, billing, and shipping information.

o Contact us with your current contact information and the information you would like to access. We will provide you the information requested if reasonably available, or will describe more fully the types of information we typically collect.

OVERSTOCK PRIVACY POLICY SCOPE

This privacy policy applies to all current or former customer information collected by or provided to Overstock.

The Sites may offer links to other sites. If you visit one of these sites, you may want to review the privacy policy on that site. In addition, you may have visited our website through a link or a banner advertisement on another site. In such cases, the site you linked from may collect information from people who click on the banner or link. You may want to refer to the privacy policies on those sites to see how they collect and use this information.

INTERNATIONAL CUSTOMER PRIVACY

In some cases, Overstock.com has partnered with companies [e.g., FiftyOne, Inc., who does business under the name "FiftyOne" and PayPal, Inc. ("PayPal"), collectively, "Vendors"] as outside vendors that we have selected to help us facilitate international transactions. We work closely with these Vendors to ensure that your transaction is handled with care and all the information you provide is secure.

As an international customer, when you click on the checkout button, you will be redirected to a checkout page hosted by FiftyOne to complete your order. On the checkout page, you will be required to select a method of payment and submit credit card and other information to FiftyOne to complete your order. On the checkout page, you will be presented with FiftyOne's terms and conditions which you must agree to in order to complete your order.

Upon completion and approval of your order by FiftyOne, FiftyOne will notify us of the approval and we will process your order and cause it to be shipped directly to FiftyOne. In this process, FiftyOne will purchase those items in your order from us, thereby taking title to the items, bill your credit card, collect and remit any duties and taxes to the appropriate taxing authority and arrange for the delivery of your order. In this process, FiftyOne makes the sale to you as the merchant of record, but we are legally obligated to deliver your order as set forth in our Terms and Conditions.

If you have questions about your order, you should direct them to us and not to FiftyOne.

The Vendors may give you the opportunity to receive marketing messages from them, in which case you should refer to their terms and conditions for details about how they use your information.

The Vendors have assured us that they will process information received from you with at least the same level of privacy protection as set forth in the US-EU Safe Harbor Framework as set forth by the US Department of Commerce regarding the collection, use, and retention of information from European Union member countries and Switzerland. If you choose to provide us and/or the Vendors with information, you consent to the transfer and storage of that information on servers located in the United States.

Any information you provide us is controlled and processed by Overstock.com, Inc., 6350 South 3000 East, Salt Lake City, Utah 84121, USA or its suppliers (where indicated herein). As mentioned herein, your information provided at checkout will be controlled and processed a Vendor. Overstock.com complies with the US-EU Safe Harbor Framework and the US-Swiss Safe Harbor Framework as established by the U.S. Department of

Commerce and approved by the European Commission and the Swiss Federal Data Protection Authority. We conduct an annual self-assessment to verify that we are in compliance with the Safe Harbor Privacy Principles in addition to our own Overstock privacy program assessments. Overstock.com has certified to the U.S. Department of Commerce and the European Union that our processing of personal data is in compliance with the Safe Harbor Privacy Principles. For more information about the Safe Harbor program, and to view our certification page, visit the Department of Commerce's Safe Harbor website. While the Safe Harbor Principles are designed to protect information originating in the European Economic area and Switzerland, our policy is to protect all international customer information in accordance with these Principles.

Customers shipping internationally who wish to inquire about or update information or change marketing preferences or anyone who wants to receive information about our international privacy program should contact us directly using one of the following methods:

Send an e-mail to customercare@overstock.com

Call our customer care line at (801) 947-3100

Write us at Overstock, 6350 South 3000 East, Salt Lake City, UT 84121, USA

In compliance with the US-EU and US-Swiss Safe Harbor Principles, Overstock.com commits to resolve complaints about your privacy and our collection or use of your information. European Union or Swiss citizens with inquiries or complaints regarding this privacy policy should first contact Overstock.com using one of the following methods:

Send an e-mail to customercare@overstock.com

Call our customer care line at (801) 947-3100

Write us at Overstock, 6350 South 3000 East, Salt Lake City, UT 84121, USA

Overstock.com has further committed to refer unresolved privacy complaints under the US-EU and US-Swiss Safe Harbor Principles to an independent dispute resolution mechanism, the BBB EU SAFE HARBOR, operated by the Council of Better Business Bureaus. If you do not receive timely acknowledgement of your complaint, or if your complaint is not satisfactorily addressed by Overstock.com, please visit the BBB EU SAFE HARBOR web site at <http://www.bbb.org/us/safe-harbor-complaints> for more information and to file a complaint.

OVERSTOCK PRIVACY POLICY REVISIONS

By interacting with Overstock, you consent to our use of information that is collected or submitted as described in this privacy policy. We may change or

add to this privacy policy, so we encourage you to review it periodically.

This Privacy Policy was last updated on January 9, 2013.

V. WASHINGTONPOST.COM

At the same time that The Washington Post Company and washingtonpost.com are committed to bringing you information tailored to your individual needs, we recognize the importance of protecting the privacy of your personally identifiable information. In adopting this privacy policy, our intent is to balance our legitimate business interests in collecting and using personally identifiable information and your reasonable expectations of privacy. Please note: this policy applies only to information collected by washingtonpost.com online, as specified below, and does not govern or apply to information collected or used by The Washington Post Company or its affiliates through other means

WHAT PERSONALLY IDENTIFIABLE INFORMATION DO I PROVIDE TO WASHINGTONPOST.COM?

Washingtonpost.com asks you to provide various types of personally identifiable information to enhance your experience on our site. During registration, washingtonpost.com asks for information such as your name, e-mail address, year of birth, gender, Zip code, country, street address, Job Title, Primary Responsibility, Job Industry and Company Size. The more information you provide, the better we are able to customize your experience. We may also ask you for other information at other times - such as when you enter a contest or participate in a promotion, when you post an online ad, when you participate in our message boards, or when you order products from us. Whenever you provide personally identifiable information to us, we will make an effort to link to our privacy policy. See below about "cookies" and what other information is collected.

HOW DOES WASHINGTONPOST.COM USE MY PERSONALLY IDENTIFIABLE INFORMATION?

Our primary goal in collecting personally identifiable information is to provide you, the user, with a customized experience on our network of sites. This includes personalization services, interactive communications, online shopping and many other types of services, most of which are completely free to you.

Washingtonpost.com uses the personally identifiable information you provide to us in several ways. Some examples follow.

A user's personally identifiable information may be used by washingtonpost.com for editorial purposes such as to contact you as part of an online survey. Additionally, we may also use the information provided by you to: 1) contact you with legal notices, 2) to advise you of any changes or additions to our Service or terms and conditions, and 3) account status (including confirmation of registrations). If you do not wish to receive the foregoing and therefore unregister from the site, please contact Customer Care and ask to have your registration account deleted. Once your account has been deleted, you will no longer have access to washingtonpost.com, however, you may reregister at any time.

We may also use information about our users and their activities on our site to send offers and information about The Washington Post Company and its affiliates. However, if you no longer wish to receive the foregoing, please contact Customer Care to request removal from this list.

Further, if you told us in your account preferences that you would be interested in receiving certain e-mail newsletters from us, we will send you those e-mail subscriptions. Please note that regardless of those subscription preferences, we may contact you by e-mail as described elsewhere in this privacy policy (e.g., to notify you about changes or additions to our Service, such as new features). However, if you no longer wish to receive any of the foregoing, you may change your preferences for the future at any time by clicking on the following link:

<https://ssl.washingtonpost.com/actmgmt/registration/addnewsletter/long>

In order to provide services free of charge, we display advertisements. washingtonpost.com delivers targeted advertisements on behalf of advertisers. Advertisers give us an advertisement and tell us the type of audience they want to reach (for example, females over 25 years old). We take the advertisement and display it to users meeting those criteria. In this process, the advertiser never has access to you washingtonpost.com account.

Washingtonpost.com also does research on our users' demographics, interests and behavior based on the information you provide to us including upon registration, on order forms, during a promotion, as well as from our server log files or from surveys. We do this to better understand and serve our users. This research is compiled and analyzed and washingtonpost.com

may share aggregated versions of this data with advertisers or other businesses. In addition, under confidentiality agreements, washingtonpost.com may match user information with third party data.

DO OTHER COMPANIES OR PEOPLE HAVE ACCESS TO PERSONALLY IDENTIFIABLE INFORMATION I PROVIDE TO WASHINGTONPOST.COM?

When you are on an area of washingtonpost.com and are asked for personally identifiable information, you are providing that information to The Washington Post Company, its divisions or affiliates, or vendors providing contractual services for washingtonpost.com (such as hosting vendors and list managers). If personally identifiable information is being provided to and/or maintained by any company other than these, our policy is that we will not transfer that personally identifiable information unless notice is given prior to transfer. If you do not want your information to be shared, you can choose not to allow the transfer by not using that particular service or by expressing this preference, if requested. Additional information about personally identifiable information follows.

Promotions: Promotions that run on washingtonpost.com may be sponsored by companies other than washingtonpost.com or may be co-sponsored by washingtonpost.com and another company. Some or all personally identifiable information provided by you during a promotion may be shared with the sponsor. If information will be shared, we will disclose such sharing prior to the transfer. You can decide not to participate in the promotion if you don't want your personally identifiable information to be shared. In certain circumstances, you may not be able to participate in a particular promotion if you chose not to share personally identifiable information. Currently, most washingtonpost.com promotions are limited to U.S. or North American residents.

Advertisers and Links: washingtonpost.com advertisers, or Web sites that have links on our site, may also collect personally identifiable information directly from you. The information practices of companies collecting data on our site or Web sites linked to washingtonpost.com are not covered by this privacy statement.

Marketplace: If you make a purchase or request a service from a business in Marketplace, the information you provide for that transaction (as well as tracking information and cookies as described below) is provided directly to the Marketplace business. Businesses listed in washingtonpost.com Marketplace have separate privacy and data collection practices.

washingtonpost.com has no responsibility or liability for these independent policies. For more information regarding the business and its privacy policy, go to that business' home page and click on the appropriate link.

Other: If we run competitions or contests on the site, you may be required to provide additional information such as your telephone number and address in order to participate. The exact rules may vary in each case but the specific rules for any contest will state how that information may be used. If you told us in your account preferences that you would be interested in receiving e-mail from us, we may send you e-mails about washingtonpost.com products, promotions, or services as well as on behalf of other companies. You can change your account preferences at any time. In addition, in each advertising email you will be provided an ability to opt-out of receiving future emails from the advertiser.

We do not control the privacy policies of our advertisers, sponsors or other sites or businesses to which we provide hyperlinks or access. Please visit the sites of these businesses to review their privacy policies.

Washingtonpost.com users should also be aware that, when you voluntarily disclose personal information in chat areas or bulletin boards, that information may be collected by others and may result in unsolicited messages from others.

Except as stated in this privacy policy or at the time of collection, a user's personally identifiable information will not be transferred to a party outside The Washington Post Company, its divisions or affiliates, or its service vendors unless notice is given at the time of collection or prior to transfer. washingtonpost.com may also disclose account information in special cases when we have reason to believe that disclosing this information may be necessary to identify, contact or bring legal action against someone who may be violating our User Agreement, may be causing injury to or interference with (either intentionally or unintentionally) washingtonpost.com's rights or property, other washingtonpost.com users, or anyone else that could be harmed by such activities, pursuant to a request from law enforcement, a subpoena, or a court order, or when otherwise may be required by law.

Data Security: We have in place physical, electronic and managerial procedures to protect the information we collect online. However, as effective as these measures are, no security system is impenetrable. We

cannot guarantee the security of our database, nor can we guarantee that the information you supply will not be intercepted while being transmitted to us over the internet.

WHAT INFORMATION DO WEB SERVERS COLLECT?

Web servers serving washingtonpost.com automatically collect certain non-personally identifiable information, such as which pages each user visits and the domain name of visitors. This information is used for various purposes including internal review, to tailor information to individual visitors and other users, for traffic audits, and as described elsewhere in this privacy policy. We also provide this information (as well as information from third-party market researchers) about our users on an aggregated, anonymous basis to our advertisers.

WHAT ARE COOKIES AND HOW DOES WASHINGTONPOST.COM USE THEM?

Washingtonpost.com places a "cookie" on the browser of a washingtonpost.com user's computer to store and sometimes track information about you. A cookie can be used to tell when your computer has contacted a Web site; we may also use the information for editorial purposes and for other purposes such as measuring certain traffic patterns. For example, cookies are used to ensure that you don't see the same ad too many times in a single session and that you do not have to reenter your login name or password during your visit. We may also use cookies to understand your use of the Service. Advertising service vendors that serve ads into our site may also use their own cookies. You may opt out of those cookies as described below. You may opt-out of the cookies delivered by washingtonpost.com by changing the setting on your browser. Please be aware that this will disable all cookies delivered to your browser, not just the ones delivered by washingtonpost.com.

HOW CAN I OPT OUT OF ONLINE ADVERTISING COOKIES?

Online advertising for washingtonpost.com is delivered by the vendor DoubleClick. DoubleClick places cookies on your browser to facilitate serving particular ads " for instance, to help determine whether you have seen a particular advertisement before, to tailor ads to you if you have visited our site before, and to avoid sending you duplicate advertisements. You can opt out of DoubleClick's use of cookies for these purposes by visiting <http://www.google.com/intl/en/privacy/>.

In some cases, we and advertisers on [washingtonpost.com](http://www.washingtonpost.com) and other sites work with other third-party vendors to help deliver advertisements tailored to your interests. These vendors include ad networks and audience segment providers, and they place cookies on your browser to collect information about your online activity (e.g., the sites and pages you have visited) in order to help advertisers deliver particular ads on our site and other sites that they believe you would find most relevant. You can opt out of those vendors' use of cookies to tailor advertising to you by visiting <http://www.aboutads.info/>.

Often our advertisers contract with a third-party service to host their ads. In this case, an ad serving vendor contacts the advertisers' hosting service for a particular advertisement. In that case, an independent cookie may be used by the third-party service. We do not have a mechanism to allow visitors to opt-out of cookies from vendors with whom we do not have a contractual relationship.

Kids under 13: Do not send any information about yourself to us - including information like your name, address or e-mail address. In general, we do not knowingly collect personally identifiable information from children under 13. If, in limited circumstances, we do knowingly collect personally identifiable information from a child under the age of 13, we will do so only with verified parental consent prior to collection. In the event that we learn that we have collected any personal information from a child under the age of 13 without verification of parental consent, we will delete that information from our database as quickly as possible.

Technology on the Internet is developing at a rapid pace, and we need to maintain our flexibility in the online arena. If we need to change our policy in the future, we will post these changes as soon as they go into effect.

© Copyright 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2010 WP Company LLC d/b/a The Washington Post

VI. WUNDERGROUND.COM

PRIVACY STATEMENT

REVISED AS OF OCTOBER 30, 2013 "About Our Privacy Statement"

The Weather Underground, LLC ("Weather Underground") is strongly committed to protecting your privacy. This Privacy Policy discloses how we

collect, use and share information we gather about you on the Services and the choices you have regarding our use of and your ability to correct the information. "Services" include Weather Underground's wunderground.com Â® website, wxquickie.com, weatherquicke.com, wundermap.com (the "Sites"), and other Internet enabled or wireless means by which Weather Underground provides content to you or receives content from you, including, without limitation, downloadable software applications (including Desktop Gadget and Dashboard Widget desktop applications), mobile web sites, mobile downloadable applications including WunderRadio, content and blog submission services, chat rooms, SMS messaging, and delivery of Weather Underground and WunderRadio content to you at your request.) "Services" also includes alert products. We hope that this disclosure will enhance your experience and reinforce the trust that you place in Weather Underground's products and services.

INFORMATION COLLECTED BY US WITHIN THE SERVICES

In order to maximize your experience and to provide you with enhanced products and services, Weather Underground may collect personally identifiable information ("PII") from you including your name, e-mail address, address, and phone number. For example, we collect PII when you become a member, sign up your weather station or webcam or weather radio, sign up for email or text alerts, sign up for our API (weather information for use on certain platforms), the Weather Desktop Gadget desktop application. Other examples of when we collect personally identifiable information are when you:

- Complete a survey
- Register for a promotion, contest or sweepstakes
- Send us an e-mail message or complete an inquiry form to receive additional information
- Submit a photo to WunderPhoto
- Report your weather conditions
- Participate in an online forum or community,
- Request tweets from us, or
- Customize the Services.

We also may collect certain information through automated means, such as:

- Specific geographic location where you are currently located
- Information about your device and device capabilities
- Information about your device operating system

Information about the applications on your device
Information about how you use the Services
Your activities on the Services
IP address
Device identifier
Carrier
Browser Type
Browser Identifier
Referring URL

Chat Sessions, Blogs, and Submissions In the case of participation in a chat session, we require that you use the handle you selected when you registered. Please be aware that chat sessions are recorded and may be monitored. Any information including personally identifiable information, that you submit during a chat session, in blog comments, in community discussions, in any other user comment field, or in connection with your content submissions, including photo submissions, can be read, collected, or used by others who access them, and could be used by other users to send you unsolicited messages. We are not responsible for the information, including personally identifiable information, you choose to submit in these forums.

Content Submissions. If you upload or rate photos in WunderPhoto or blog comments, you have to be registered and logged into this feature, and you will be required to provide us with a handle and email address. If you provide us with content, including, without limitation, submissions to WunderPhoto and user comments, any metadata, including, without limitation, geo-location and tags, in that content will become publicly available.

Location Information. In certain instances in our Services, you will be able to input specific location information in order for us to bring up the weather for that location. For that Service, we will ask you to input specific location such as address, city and state, latitude and longitude, country, neighborhood name, beach location, or zip code.

To offer you certain Services, provide advertisements based on your physical geographic location, and engage in the other uses described below, on some devices we provide you the option to opt-in to allow us to use your physical geographic location using GPS, cellular network location based services on your device, or browser services. You may at any time opt-out from these geo-location services by going to the settings on your mobile device.

If you have the WunderRadio mobile application installed on your Microsoft Windows 7 device, you may also opt-out from geo-location by going to the "Settings" menu in the application and toggling location services off. Toggling location services off may also disable certain services in the application.

Alerts and notifications. In certain of our Services, you have the ability to receive push notifications for weather related updates. To provide these services, we may need to collect your email address, phone number, device information, and mobile carrier information, in addition to the zip code or geographic area to which the alert, update or notification pertains. If you no longer wish to receive these types of notifications you may opt-out by turning them off within the settings of your mobile device or the particular Service for which you registered.

HOW WE USE AND SHARE YOUR INFORMATION

One of the most valuable assets of our business is you. Except as described below, Weather Underground does not rent, sell or share personally identifiable information about you with other people or nonaffiliated companies.

A. Weather Underground and its Affiliates

We use information collected on the Services, and disclose information to third parties, including your physical geographic location, to help fulfill your requests or in connection with the operation of the Services, for example to service your account, conduct a transaction or provide a service that you have requested, display information and advertisements that we believe match your interests and profile, facilitate updates to the Services, provide support, notify you if you've won a contest or sweepstakes, improve the working of the Services, and compile statistics about our users and your use of the Services. We and such third parties may on occasion combine information about you with information obtained from other parties to market to you products or services that may be of interest to you. If you sign up for communications from Weather Underground, they may also contain offers from third parties.

At various points, we may provide you with the opportunity to opt in to receive special offers from Weather Underground and its affiliates. If you choose to opt in, Weather Underground and its affiliates will use your personally identifiable information to inform you of other products or

services available from Weather Underground and its affiliates, including The Weather Channel, LLC, WSI Corporation and Weather Central, LP. These products may include, for example, schedules for special programs on The Weather Channel cable network or offers to purchase a calendar from The Weather Channel.

B. Companies Offering Promotions, Products, or Services

At times we will provide you with the opportunity to opt-in to receive promotions, products or services from third party sponsors. With your permission, Weather Underground will periodically send to you, via e-mail, offers and promotions from our third party sponsors which we select based upon your interests or user profile. In this case we do not share your information with the third party sponsors—we do the mailing for them. You will always have the opportunity to opt-out of receiving future offers by following the unsubscribe link included with each email. In other instances, we provide you with the opportunity to receive products and services directly from specifically identified third parties through the Services. In these instances, you must explicitly consent to receive each offer. When you explicitly agree to receive offers from specific third parties, your information is shared with the specific third party to enable them to provide you with the offer you have requested. We are not responsible for the information collection practices of these third parties, and all information provided is governed by the privacy policies of these third parties. You should review the privacy policies of these parties before supplying personally identifiable information to them.

C. Third-Party Processors

When your information is collected on the Services, it may be collected directly by or shared with a selected third parties in connection with the operation of the Services or the provision of services to you ("Third Party Processors"). These Third Party Processors may include, for example, companies that operate our online WunderStore, that process credit card information or handle shipping for Weather Underground, that deliver materials to you via e-mail or postal service, that organize, administer, process, or provide advertising services, and/ or that analyze data on our behalf to help us provide more relevant offers to you and to eliminate the delivery of duplicate offers as well as correcting and/or updating users' data based on information we provide them. We also provide your information to third party mapping service providers to provide you map content for the

Services. Third party processors are authorized to use your information that we provide to them only to carry out the service they are providing for Weather Underground. We require that the third party processors securely store the personally identifiable information we provide to them and maintain it so that it can be accessed when needed for future purposes, and all such Third Party Processors are contractually bound by us to keep the personally identifiable information confidential.

D. Other Web Sites, Other Services, and Links

Like many Web sites on the Internet and other Internet-based services, many of our Services link to pages located on Web sites or services maintained by various other entities. In some cases you may link to pages of other Web sites that are framed with elements of our Services such as in the header or footer. In that case the URL will identify the site you are visiting. In other cases, such as advertisements, you will be connecting to another site or service when you click on or otherwise activate those opportunities, including click-to-call, click-to-text, and click-to-email opportunities. These other sites and services are not bound by our privacy policy, and we are not responsible for their information collection practices. The privacy policies of other organizations may differ from ours, for example, with respect to the level of security, use of cookies, and collection, use and disclosure of personally identifiable information.

Some of our Services allow users to interface with other sites or services, including Facebook and Twitter. You will remain logged in to those other sites and services until you log off those sites and services. By proceeding with these interfaces, you are allowing our Services to access your information on those other sites and services, and you are agreeing to those other sites' and services' Terms of Use. Once you log into those other sites and services, the content you post on those other sites and services may also post to our Services. Your use of those other sites and services is subject to the privacy policies of those sites and services, and not this Privacy Policy.

Some of our Services use third party operating systems, platforms, communication services, devices, and software elements (such as mobile device operating systems, wireless services, mobile phone and tablet devices), and some of our Services are provided by third party distributors, device makers, device operators, platform operators, and communication services. Weather Underground does not control these third party entities, products and services, and they may collect, use, process, transmit and disclose your

information. For example, these third party elements may gather information from our Services, including our applications, in a manner that overrides your settings and preferences in our Services. As Weather Underground does not control or have knowledge of their data handling practices, we recommend that you review their privacy notices, terms of use and license agreements.

E. Use by us of IP Addresses, Cookies, Web Beacons, Single-pixel Gifs, and Information Saved by Other Technologies

Weather Underground and Third Party Processors acting on our behalf use various Internet technologies which help us to manage the operations of our Services and track usage behavior so that we can tailor information and advertisements to make your visits more enjoyable and meaningful. These Internet technologies include:

IP Addresses: numbers assigned to individual computers which accompany every packet of information across the Internet.

Cookies: small electronic files transferred from our services to your electronic device.

Beacons and single-pixel Gifs: electronic tags.

HTML5 LocalStorage: storage within your browser on your electronic device.

Information that Weather Underground collects via the technologies described in this section may be linked to other information about you. Weather Underground uses the technologies described in this section, alone or in combination, to provide and administer our services, understand user behavior, manage services operations, measure the effectiveness of advertisements and our service operations, target advertising, help diagnose problems, recognize repeat visitors, track and analyze usage behavior, facilitate your access to and use of our services, serve, target, control, and measure advertisements on our Services, count users who have visited certain areas of our Services we wish to track, and help determine the effectiveness of advertising campaigns on our services. In addition, we may share demographic information, location data, IP address, aggregate (not individual) usage statistics for our Services, other identifiers and information with advertisers and other third parties. For example, we may share IP address, random or anonymous device identifier, city and state, ZIP code, and specific geo-location with the parties identified in subparagraph F below.

Your browser may allow you to manage your cookies and HTML5 LocalStorage. Each browser is a little different, so look at your browser Help menu to learn the correct way to modify your cookies and HTML5 LocalStorage. If you turn Cookies or HTML5 Local Storage off, however, you won't have access to many of the features which make your experience more efficient and some of our Services will not function properly.

Weather Underground contracts with certain Third Party Processors to track, analyze and report data about the usage of our Services using these Internet technologies. These Third Party Processors include Google Analytics, Paypal, and Omniture.

F. Use of Cookie, Web Beacon and Other Technologies by Advertisers, Ad Networks, Ad Servers and Analytics Companies

Weather Underground works with a variety of advertisers, advertising networks, advertising servers, and analytics companies. These advertisers, advertising networks, advertising servers, and analytics companies use various technologies, and technologies from third party companies, to collect data in order to send (or serve) relevant ads to users. These technologies may include the placement of cookies or web beacons, the use of HTML5 LocalStorage, the use of unique or non-unique non-personal identifiers, or the use of other technologies on our Services, and these technologies may be used to track user behavior, to track how our Services are being used, and possibly to serve you more relevant ads. These targeted advertisements may appear on our Services or other services that you visit. This Privacy policy does not cover the use of various technologies by advertisers, advertising networks, advertising servers, and analytics companies. These companies may also obtain information from services you use from other companies, including without limitation, other websites, mobile website, mobile downloadable applications, and downloadable applications, and combine that information with information they obtain through these third party technologies on our Services. You should be aware that The Weather Underground does not have control over these third party technologies or the information contained in them. For more information about how DoubleClick, and some of the other ad networks, ad servers, analytics companies, and third party companies use the information collected by the technologies on our services and about your option not to accept cookies placed by some of these companies on our Services, please visit the following sites:

DoubleClick: http://doubleclick.net/privacy_policy,

Nielsen: <http://www.nielsen-netratings.com/privacy/sitecensus.htm>

Quantcast: <https://www.quantcast.com/how-we-do-it/consumer-choice/privacy-policy/>

Flurry: <http://www.flurry.com/privacy-policy.html>

Google Analytics: <http://www.google.com/analytics/learn/privacy.html>

Omniure: <http://www.omniure.com/static/61>

comScore and its affiliate, Scorecard Research:

<http://www.scorecardresearch.com/Priv.html>

http://networkadvertising.org/managing/opt_out.asp

and <http://www.aboutads.info/choices>

These opt-outs are device specific and may not work on all devices. You should be aware that if you choose to opt out through any of these sites, this does not opt you out of being served advertising. The ads will just not be targeted to you by any party from which you have opted out.

You can also opt out of future information collection from our Services by ceasing use of the Service or in the case of an application, uninstalling the application.

G. Purchase or Sale of Businesses

Weather Underground continually looks for ways to improve our business, including purchasing or selling all or part of a business or company. If we buy or sell a business, personally identifiable information and non-personal information will likely be transferred as part of the sale. If we buy a business, we will honor the requests of that business' former customers regarding their personally identifiable information. If in connection with a sale there will be material changes to the way personally identifiable information is being used, affected consumers who have provided us with an email address will be notified via e-mail, if you have provided us an email address, and will be given the opportunity to opt-out.

H. Other Limitations on Privacy

We may provide personally identifiable information without your permission (i) pursuant to judicial or other government subpoenas, warrants, or orders or otherwise to comply with law; (ii) where The Weather Underground believes the rights, property or an individual's safety or security is at risk; (iii) if we find that your actions violate our Terms and Conditions of Use; or (iv) where otherwise required by law.

E-MAIL A FRIEND

If you choose to use our service that allows you to e-mail parts of our Services to a friend, in most circumstances we need to have access to your and your friend's e-mail addresses temporarily to send the e-mail due to technical requirements. This information is not used for any other purpose.

MOBILE DEVICE CONTACTS

Some mobile products allow you to view weather conditions for people in your phone's contact list. If you agree to use this feature, your contact list will be accessed in order for us to provide weather information for your contacts who already have their address, city, state, or zip code available. We do not store this information or use it for any other purpose.

OPTING-OUT, CORRECTIONS AND CANCELLATIONS

If you receive an offer or promotion from Weather Underground or a third party related to Weather Underground, you must have opted in. To ensure we have your consent, we make sure that each e-mail sent to you by Weather Underground presents the additional opportunity to opt-out of receiving future e-mails.

If you have registered with us and have obtained a password, you may reset your password, update your personally identifiable information, correct information, make changes to your preferences, unsubscribe to any e-mail subscriptions or opt-out of certain services, by accessing your profile online at the member settings page or within the settings of a particular application or you may contact us directly at wuhelp@wunderground.com.

DATA RETENTION OF ACCOUNT INFORMATION

If you wish to cancel your account or request that we delete or no longer use your account information to provide you Services, contact us at wuhelp@wunderground.com. Subject to applicable law, we will retain and use your account information as necessary to comply with our legal obligations, resolve disputes, and enforce our agreements.

CHILDREN'S PRIVACY

Weather Underground is committed to protecting the safety and privacy of young people using the Internet. We do not knowingly collect personally identifiable information from children under the age of 13 and believe that children should get their parents' or guardians' consent before giving out any personally identifiable information.

The Services may contain links to third party Web sites and other services geared to children. The linked sites and services are not under the control of Weather Underground. We are not responsible for the contents of any linked site or service or any link contained in a linked site or service, nor are we responsible for the privacy practices of the operators of those sites or services with respect to children. We strongly encourage parents and guardians to review the privacy policies of all linked sites and services.

INTERNATIONAL TRANSFER

Weather Underground, and all associated Services and systems, including registration, is housed on servers operated in the United States. If you are located in the European Union or elsewhere outside of the United States, please be aware that information we collect (including cookies) will be processed and stored in the United States, a jurisdiction in which the data protection and privacy laws may not offer the same level of protection as those in the country where you reside or are a citizen. By using our Services and providing information to use, you consent to the transfer to and processing of the information in the United States.

OUR COMMITMENT TO SECURITY

Keeping the information you provide to Weather Underground secure is important to us. We have put in place appropriate physical, electronic and managerial procedures designed to protect and safeguard your data. Of course, although Weather Underground uses standard industry practices to protect your personally identifiable information, we cannot guarantee that your communications with the Services will never be unlawfully intercepted, or that your personally identifiable information will never be unlawfully accessed by third parties.

CHANGES TO THE PRIVACY POLICY

Weather Underground reserves the right to alter our Privacy Policy as business needs require. If we decide to change our Privacy Policy, we will post those changes here so that you will always know what information we gather, how we might use that information and whether we will disclose it to anyone. All changes to this policy will be posted on our Web site prior to the time they take effect. In the event that we make material changes to the way we use personally identifiable information, affected consumers who have provided us with an e-mail address will be notified via e-mail and will be given the opportunity to opt-out.

The Weather Underground, LLC

P.O. Box 724554

Atlanta, GA 31139

DATA FEEDS

If you connect to Wunderground Data Feed, you will be able to pull and to decode weather related data from the non-public website of Weather Underground, Inc. ("WUI").

Your access to Wunderground Data Feed, and your use of the data pulled from Wunderground Data Feed, is subject to the Terms of Use generally applicable to all features of www.wunderground.com. (the "Site")

In addition, because of the special features of Wunderground Data Feed, by accessing and viewing information from Wunderground Data Feed you agree to be bound by these special terms: (1) you will not modify the data from Wunderground Data Feed in any way; (2) you will not modify the link structure to Wunderground Data Feed; (3) you will not modify WUI's logo that is included in the data; and (4) in all uses of the data, you will credit WUI by name as the source of the data.

As with all data that you access from the Site, the data that you pull from Wunderground Data Feed may only be used by you for personal, non-commercial purposes. If you want to use that data for commercial purposes, contact us through the Site and we will supply you with rate information for commercial customers.

If you violate any of the general or specific terms applicable to Wunderground Data Feed, WUI has the right to terminate your use of the Site and the Data Feed and to take appropriate legal actions against you.

WUI reserves the right to change these general or special terms at any time by posting on the Site. You understand and agree that your use of the Site and Wunderground Data Feed is a benefit voluntarily given by WUI and that WUI may withdraw that benefit and rescind your participation at any time for any reason in its sole discretion. If you do not agree to these general and special terms, you cannot use Wunderground Data Feed or the services and information offered therein.

UNDERSTANDING THE FEDERAL CIRCUIT: AN EXPERT COMMUNITY APPROACH

Laura G. Pedraza-Fariña[†]

ABSTRACT

The Court of Appeals for the Federal Circuit (“CAFC”)—the appeals court in charge of virtually all patent cases—has been fraught with controversy since its creation in 1982. To its critics, the Federal Circuit engages in puzzling behaviors, out of step with its role as an Article III appellate court. The Federal Circuit shows little deference to district courts on questions of fact and to the Patent and Trademark Office (“PTO”) on technical issues. It surprisingly resorts to formalistic rules in an area of the law that requires flexibility to adapt to changing technological landscapes. These criticisms have become increasingly salient, leading to calls for an end to the Federal Circuit’s exclusive jurisdiction over patent appeals. Several explanations have been put forth to account for these puzzling behaviors. Yet, none can fully explain the range of unique Federal Circuit conduct. Without a full explanation for Federal Circuit behavior, however, the debate over Federal Circuit jurisdiction will remain gridlocked.

Drawing upon studies from the sociology of expertise, this Article provides a model of Federal Circuit decision-making that explains and predicts Federal Circuit behavior as a product of four distinct but interrelated expert community features: (1) epistemic control, (2) codification, (3) typecasting, and (4) inability to self-coordinate. The drive that expert communities exhibit for maximal control and autonomy over their knowledge base—referred to as epistemic control—explains why the Federal Circuit is less likely to defer to solutions proposed by other expert communities, such as the PTO, than would be expected of generalist courts. Those motivations also predict that expert communities such as the Federal Circuit will be more likely to defy non-expert superior generalists, such as the Supreme Court, than would be expected under traditional accounts of behavior in judicial hierarchies. The codification feature of expert communities gives a richer account than

© 2015 Laura G. Pedraza-Fariña

[†] Assistant Professor, Northwestern University Law School, laura.pedraza-farina@law.northwestern.edu. I thank Jonas Anderson, Adam Candeub, Brian Casey, Kevin Emerson Collins, Mariana Craciun, Shari Diamond, Rochelle Dreyfuss, Peter diCola, Gil Eyal, Roger Ford, Ezra Friedman, John Golden, Paul Gugliuzza, Shubha Ghosh, Tonja Jacobi, Andrew Koppelman, Joshua Kleinfeld, Peter Lee, Orly Lobel, Peter Menell, Jide Nzelibe, Lisa Larrimore Ouellette, Matthew Sag, Joshua Sarnoff, David Schwartz, Nadav Shoked, Matthew Spitzer, Kristen Stilt, Melissa Wasserman, and audiences at Northwestern Law School, the American Bar Foundation, DePaul Law School, the 2014 Intellectual Property Scholars Conference, PatCon 2014, Works-in-Progress in Intellectual Property 2014, the 7th Annual Junior Scholars in Intellectual Property Conference, Michigan State University Law School, and the 2014 Intellectual Property Scholars Roundtable for their helpful comments and feedback. Jason Portis and Ryan Whalen provided excellent research assistance.

existing narratives of when and why the Federal Circuit may prefer inflexible rules of decision over flexible standards. It predicts that the Federal Circuit will resort to rules not only to simplify technical knowledge or control subordinate communities, but also to build external legitimacy and manage internal dissent. Normatively, this model offers a path out of the gridlock by revealing a framework to evaluate and design proposals for Federal Circuit reform. To minimize the distortive effects of typecasting in the context of a centralized court, while retaining the advantages of expertise, this Article proposes the use of advisory panels to house technological, sociological, and economic expertise. Additionally, the model has important implications beyond the Federal Circuit, as it provides a novel theoretical lens to analyze the behavior of other specialized courts.

TABLE OF CONTENTS

I.	INTRODUCTION	91
II.	THE FEDERAL CIRCUIT UNDER FIRE	97
A.	RELATIONSHIP WITH LOWER COURTS AND AGENCIES	98
1.	<i>Relationship with Lower Courts</i>	98
2.	<i>Relationship with the Patent and Trademark Office and the International Trade Commission</i>	102
B.	ADJUDICATIVE “FORM” OF PATENT LAW.....	105
III.	APPROACHES TO EXPERT DECISION-MAKING	108
A.	THE INSTITUTIONAL APPROACH TO EXPERTISE: SOCIOLOGY OF THE PROFESSIONS	108
1.	<i>Competing Framings of Tasks and Problems</i>	110
2.	<i>Rule-Making or Codification</i>	111
B.	THE EMBODIED APPROACH TO EXPERTISE: HOW NOVICES BECOME EXPERTS.....	114
IV.	A TYPOLOGY OF FEATURES OF EXPERT COMMUNITIES	116
A.	THE FEDERAL CIRCUIT AS AN EXPERT COMMUNITY	116
B.	FOUR FEATURES OF EXPERT COMMUNITIES AND THEIR APPLICATION TO FEDERAL CIRCUIT BEHAVIOR	120
1.	<i>Epistemic Control</i>	120
a)	Explaining the Federal Circuit’s Relationship with Lower Courts: Non-Deference	121
b)	Explaining the Federal Circuit’s Relationship with Other Expert Bodies: Jurisdictional Competition	123
c)	Predicting Federal Circuit Behavior: Defying the Supreme Court	123
d)	Predicting Federal Circuit Behavior: Jurisdictional Expansion	127
2.	<i>Codification</i>	128

a)	Explaining the Federal Circuit’s Relationship with Lower Courts and Agencies: Teaching, Delegation, and Control.....	129
b)	Predicting External Relationships: Seeking Legitimacy through Rules	132
c)	Predicting Internal Relationships: Managing Dissent	134
3.	<i>Typecasting</i>	136
a)	Professional Typecasting in <i>Diamond v. Chakerabarty</i>	137
b)	Technical Typecasting in <i>Myriad Genetics</i>	138
4.	<i>Inability to self-coordinate across multiple expert areas</i>	140
C.	ALTERNATIVE EXPLANATIONS OF FEDERAL CIRCUIT BEHAVIOR.....	143
1.	<i>Centralization Theory</i>	143
2.	<i>Information-Costs Theory</i>	145
3.	<i>Capture and Tunnel Vision</i>	146
V.	NORMATIVE IMPLICATIONS: EVALUATING PROPOSALS FOR FEDERAL CIRCUIT REFORM	149
A.	THE DANGERS OF TYPECASTING IN A CENTRALIZED EXPERT COURT.....	151
B.	OVERCOMING COORDINATION DIFFICULTIES.....	153
C.	EXTERNAL ADVISORY PANELS AS A PROPOSED SOLUTION TO TYPECASTING AND COORDINATION PROBLEMS.....	154
VI.	CONCLUSION	159
I.	INTRODUCTION	

Patent law has transitioned from an arcane topic¹ to a field that is increasingly forced to confront some of the thorniest issues of national public policy, such as the patentability of genes,² diagnostic methods,³ and

1. See, e.g., ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS 9 (2004) (arguing that the Supreme Court rarely heard patent cases before 1982 because “[t]he justices were reluctant to devote their time to these ‘banal’ commercial disputes”).

2. See, e.g., *Ass’n for Molecular Pathology v. Myriad Genetics, Inc.*, 133 S. Ct. 2107, 2111 (2013) (holding that isolated genomic DNA is patent ineligible).

3. See, e.g., *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 132 S. Ct. 1289, 1294 (2012) (holding that a diagnostic method that helped doctors determine the appropriate dose of a drug was an unpatentable law of nature).

synthetic biology.⁴ As patent law captures national headlines,⁵ commentators have placed renewed focus on the workings of the U.S. Court of Appeals for the Federal Circuit (“CAFC” or the “Federal Circuit”)—the appellate court for virtually all patent cases.⁶

The Federal Circuit’s patent jurisprudence has come under sustained criticism. Commentators rue the Federal Circuit’s increasing preference for simple rules over standards,⁷ its unwillingness to defer to district court and the Patent and Trademark Office (“PTO”) findings of fact⁸ or to determinations by the International Trade Commission (“ITC”),⁹ its propensity for de novo review,¹⁰ and its overly expansive view of its own

4. See generally Sapna Kumar & Arti Rai, *Synthetic Biology: The Intellectual Property Puzzle*, 85 TEX. L. REV. 1745 (2007).

5. See, e.g., Charles Duhigg & Steve Lohr, *The Patent, Used as a Sword*, N.Y. TIMES (Oct. 7, 2012), <http://www.nytimes.com/2012/10/08/technology/patent-wars-among-tech-giants-can-stifle-competition.html> (discussing offensive patent litigation among technology companies); Editorial, *Clarifying, and Tightening, Patent Law*, N.Y. TIMES (June 4, 2014), <http://www.nytimes.com/2014/06/05/opinion/clarifying-and-tightening-patent-law.html> (discussing recent Supreme Court opinions on patent law); *Take Back Your Genes*, AM. CIVIL LIBERTIES UNION, <https://www.aclu.org/take-back-your-genes/> (last visited July 8, 2014) (reporting on the Supreme Court decision in *Myriad*).

6. See, for example, the remarks of Hon. Diane Wood, Chief Judge of the 7th Circuit, in which she argued that the Federal Circuit’s exclusive jurisdiction over patent cases should be abolished. *Keynote Address: Is It Time to Abolish the Federal Circuit’s Exclusive Jurisdiction in Patent Cases?*, 13 CHI.-KENT J. INTELL. PROP. 1 (2013). For a general recent discussion of Federal Circuit jurisdiction see the responses to Hon. Diane Wood’s keynote in 13 CHI.-KENT J. INTELL. PROP. (2014).

7. See, e.g., John R. Thomas, *Formalism at the Federal Circuit*, 52 AM. U. L. REV. 771, 777 (2003); Rochelle Cooper Dreyfuss, *In Search of Institutional Identity: The Federal Circuit comes of Age*, 23 BERKELEY TECH. L. J. 787, 802-04 (2008); Timothy R. Holbrook, *Substantive Versus Process-Based Formalism in Claim Construction*, 9 LEWIS & CLARK L. REV. 123 (2005); Peter Lee, *Patent Law and the Two Cultures*, 120 YALE L.J. 2 (2010); Lucas S. Osborn, *Instrumentalism at the Federal Circuit*, 56 ST. LOUIS U. L.J. 419 (2012); but cf. David O. Taylor, *Formalism and Antiformalism in Patent Law Adjudication: Rules and Standards*, 46 CONN. L. REV. 415, 486 (2013) (arguing that the Federal Circuit is justified in adopting rule-formalism if it deems that the need for certainty in a particular area of patent law outweighs flexibility and accuracy concerns).

8. See, e.g., Ted D. Lee & Michelle Evans, *The Charade: Trying a Patent Case to All “Three” Juries*, 8 TEX. INTELL. PROP. L.J. 1, 14 (1999); Arti K. Rai, *Engaging Facts and Policy: A Multi-Institutional Approach to Patent System Reform*, 103 COLUM. L. REV. 1035 (2003) [hereinafter Rai, *Engaging Facts and Policy*]; Arti Rai, *Addressing the Patent Gold Rush: The Role of Deference to PTO Patent Denials*, 2 WASH. U. J.L. & POL’Y 199, 201 (2000).

9. See, e.g., Sapna Kumar, *Expert Court, Expert Agency*, 44 U.C. DAVIS L. REV. 1547, 1550–51 (2011); Rai, *Engaging Facts and Policy*, *supra* note 8, at 1052–56.

10. See, e.g., Rai, *Engaging Facts and Policy*, *supra* note 8, at 1057–65; Rochelle Cooper Dreyfuss, *Giving the Federal Circuit a Run for its Money: Challenging Patents in the PTAB*, NOTRE DAME L. REV. 28-33 (forthcoming 2015), available at <http://ssrn.com/abstract=2572647>.

jurisdiction.¹¹ Critics argue that, in the aggregate, these features have given rise to a court that is unresponsive to the needs of communities of innovators and out of step with national innovation policy.¹² For example, John Thomas links what he terms the Federal Circuit's "adjudicative rule formalism" to the court's inability to adapt patent law to the changing conditions of technological innovation.¹³ Arti Rai characterizes the Court's aggressive de novo review as having "problematic effects across entire fields of technology."¹⁴ Sapna Kumar criticizes the Federal Circuit's unwillingness to defer to the ITC's patent decisions, in view of the ITC's greater political accountability and fact-finding capability.¹⁵ And Rochelle Dreyfuss cautions that the Federal Circuit's insistence on "particular analytical frameworks," and experimentation with "rigid interpretative methodologies," has made the law more precise at the expense of quality and accuracy.¹⁶ In part echoing these critiques, Chief Judge Diane Wood recently proposed that we eliminate the Federal Circuit's exclusive jurisdiction over patent law.¹⁷

These critiques are particularly troubling, as they flatly contradict a crucial assumption underlying the creation of the Federal Circuit: that placing the "unusually complex [and] technically difficult" patent cases in the hands of a single appeals court would lead not only to national uniformity but also to better-quality patent decisions.¹⁸ Thus, explaining these particular features of Federal Circuit jurisprudence is important both to provide a diagnosis of the current "patent failure"¹⁹ and to design a way out.

Drawing from and expanding upon studies in the sociology of expertise, this Article provides a novel perspective on the behavior of the Federal Circuit. Its central thesis is that understanding the Federal Circuit and improving judicial decision-making in patent law requires understanding expert communities, including their typical blind spots and strengths, and solutions that have been adopted in other fields to optimize centralized

11. See, e.g., Paul R. Gugliuzza, *Patent Law Federalism*, 2014 WIS. L. REV. 11, 30 (2014).

12. See, e.g., Rai, *Engaging Facts and Policy*, *supra* note 8, at 1065; Thomas, *supra* note 7, at 775; Rochelle Cooper Dreyfuss, *Percolation, Uniformity, and Coherent Adjudication*, 66 S.M.U.L. REV. 505, 530 (2013) ("The core frustration [with the Federal Circuit] is "with rigid rules that are not adaptable to new technologies or to new business models and do not take account of the public's interest in access.").

13. Thomas, *supra* note 7, at 775.

14. Rai, *Engaging Facts and Policy*, *supra* note 8, at 1065.

15. Kumar, *supra* note 9, at 1552–53.

16. Dreyfuss, *supra* note 7, at 802–03.

17. Wood, *supra* note 6, at 9.

18. H.R. Rep. No. 97-312, at 22–23 (1981).

19. See JAMES BESSEN & MICHAEL J. MEURER, *PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK* (2008).

expert decision-making. An “expert community,” as the term is used here, refers to an institutionalized group of experts that develop and apply a system of abstract knowledge to address a specific set of questions. For example, psychologists have developed a system of abstract knowledge—partially codified in the Diagnostic and Statistical Manual of Mental Disorders (“DSM”)—to diagnose and treat mental illness.²⁰ Similarly, the Federal Circuit has developed its own system of abstract knowledge—a system to identify those inventions that require a patent to incentivize innovation, and designed to ultimately foster the “progress of science and useful arts.”²¹ I develop the concept of “expert community” as it relates to the Federal Circuit throughout this article. A key claim is that there are important, but unexamined, differences between how expert and non-expert generalists will decide cases and interact with other relevant actors, and in particular with other institutional actors such as agencies, district courts, other appellate courts, and the Supreme Court. Appreciating these differences is crucial to understanding the behavior of the Federal Circuit, and of specialized courts more broadly.

Sociologists have long been interested in understanding the development, organization, and control of expertise in society. This Article draws from two lines of sociological research. The first, the institutional approach to expertise, explores how expertise is institutionalized in organized groups of experts and how those organized groups interact with each other and with society at large.²² The second, the embodied approach to expertise, focuses on understanding the transition from novice to expert and, more specifically, how decision-making processes change once novices acquire expertise in a particular domain.²³

Taken together, four basic insights emerge from these studies. First, organized groups of experts seek maximal control in the development and application of the abstract knowledge base that constitutes their expertise (what this Article analyzes as the *epistemic control* feature of expert communities).²⁴ But, expert groups are embedded in an ecosystem composed of other expert groups with different knowledge systems that apply to overlapping sets of problems.²⁵ This overlap leads expert communities to

20. See AM. PSYCHIATRIC ASS'N, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS (5th ed. 2013).

21. U.S. CONST. art. I, § 8, cl. 8.

22. See *infra* Section III.A

23. See *infra* Section III.B

24. See *infra* Subsection IV.B.1.

25. See *id.*

engage in constant competition with one another for jurisdictional control.²⁶ As applied to the Federal Circuit, this unappreciated dynamic of jurisdictional competition between expert communities can explain the Federal Circuit's rigorous, non-deferential standard of review of PTO decisions—a behavior that stands in sharp contrast with the behavior of non-expert appellate courts with respect to findings of fact by administrative agencies.²⁷ In addition, the tendency for expert communities to struggle to maintain autonomy in the development of their respective knowledge bases predicts that expert communities will be more likely to defy solutions imposed by non-expert generalists than communities of non-experts.²⁸ In the specific case considered here, the lens of epistemic control predicts that the Federal Circuit is more likely to defy Supreme Court decisions than are other circuit courts.²⁹

Second, these studies present a more nuanced picture than existing accounts of why and when an expert community, such as the Federal Circuit, will prefer formal rules to flexible standards (what this Article analyzes as the *codification* feature of expert communities).³⁰ Expert practitioners differ from novices in their relationship to the use of rules to solve problems.³¹ Novices must self-consciously follow explicit rules to begin their path towards expertise—the teaching function of rules.³² In contrast, experts can draw on a wealth of contextual information gathered through training and practice that is not readily reduced to a set of written rules of decision (or what is often termed “tacit knowledge”).³³ A direct consequence of expert tacit knowledge is an unavoidable conflict between the rules as explained to novices and their actual application by experts to real-world conflicts. But rules serve three additional functions. They lend legitimacy to expert communities by showing non-expert novices and the public at large the

26. *See infra* Subsection IV.B.1.b).

27. *See infra* Subsection IV.B.1.a). *See also* Banks Miller & Brett Curry, *Experts Judging Experts: The Role of Expertise in Reviewing Agency Decision Making*, 38 LAW & SOC. INQUIRY 55, 55–56 (2013) (noting that courts are generally deferential to the decision-making of federal agencies).

28. *See infra* Subsection IV.A

29. There have been no quantitative empirical studies comparing Federal Circuit defiance of Supreme Court decisions to defiance by other circuits, or assessing whether the Federal Circuit is more likely to defy the Supreme Court in its perceived area of expertise (patent law) than in any of the other cases that make up its docket—both of which are predicted by this model. But qualitative evidence suggests this is indeed the case. *See infra* Part IV.B.1.c).

30. *See infra* Subsection IV.B.2.

31. *See infra* Subsection IV.B.2.a).

32. *See infra* Subsection IV.B.2.b).

33. *See infra* Subsection IV.B.2.a).

utility of the expert practice.³⁴ Rules can also help manage internal dissent by reducing variability in highly fractured expert communities.³⁵ Finally, because rules prevent recourse to more subjective contextual judgment, they are also an important mechanism of jurisdictional competition: expert communities often resort to rules to constrain and control the action of subordinate communities.³⁶ As applied to the Federal Circuit, the codification feature of expert communities explains the Court's preference for rules over standards as a mechanism to: (1) constrain subordinate expert communities, such as the PTO, (2) both teach and constrain the actions of district courts (conceptualized as subordinate generalist communities), (3) legitimize Federal Circuit expertise in the eyes of relevant audiences (such as the patent bar, scientists, academics, and the Supreme Court), and (4) manage internal dissent.

Third, expert communities compete for jurisdictional control by framing overlapping problems as best solved through a particular expert community's own abstract knowledge system (what this Article analyzes as the *typcasting* feature of expert communities).³⁷ Typecasting, however, has a clear shortcoming: when expert communities typecast a particular problem as similar to other problems already solved within their domain of expertise, they are less likely to look widely for other—potentially better—available solutions.

Fourth, studies of embodied expertise teach us that as novices become experts, they also grow progressively more emotionally attached to their area of expertise. This means that issues within an expert's area of expertise are particularly personally salient to that expert relative to the general public, and relative to other issues in other fields of expertise. When the best approach to a problem requires cooperation and coordination between two or more expert areas, a community with expertise in one area is likely to place inadequate weight on the competing considerations and interests of other expert communities (what this Article analyzes as the *inability to self-coordinate* feature of expert communities).³⁸

The remainder of the Article proceeds in three parts. Part II offers an overview of the critiques surrounding the performance of the Federal Circuit

34. Because rules articulate the steps taken by an expert to solve a particular problem, they help make explicit, and thus more transparent, the mechanisms of expert decision-making. *See infra* Subsection III.A.2.

35. *See infra* Subsection IV.B.2.c).

36. *See infra* Subsection IV.B.2.a).

37. *See infra* Subsection IV.B.3.

38. *See infra* Subsection IV.B.4.

as the sole appellate arbiter of patent disputes. Part III introduces the reader to studies in the sociology of expertise. Drawing and expanding upon these studies, Part IV develops a typology of four features that are closely associated with communities with attributed expertise in a particular subject matter: (1) epistemological control, (2) codification, (3) typecasting, and (4) inability to self-coordinate. This part demonstrates how conceptualizing the Federal Circuit as an expert community helps explain key features of its jurisprudence and predict additional behaviors. Part V develops the normative implications of an “expert community” analysis of the Federal Circuit. It argues that two features of expert communities—typecasting and inability to self-coordinate—are undesirable in a centralized specialized court. To minimize the distortive effects of typecasting in the context of a centralized court, this Article proposes the use of advisory panels to house technological, sociological, and economic expertise, a strategy that is widely used to optimize medical decision-making.

II. THE FEDERAL CIRCUIT UNDER FIRE

The Federal Circuit stands alone as the only Article III court with virtually exclusive appellate jurisdiction over a specific subject matter—patent law.³⁹ Despite generally favorable assessments of its performance during the first five years of its existence since its formation in 1982, criticism of the Federal Circuit began to mount in the early 1990s and has continued to this date.⁴⁰ Indeed, both academic commentators and judges have renewed calls to abolish the Federal Circuit’s exclusive jurisdiction over patent appeals—judging this specialized court to be the cause of an ossified jurisprudence that is out of step with the needs of communities of innovators.⁴¹ This Part places this Article’s contribution in the context of current debates surrounding the institutional design of the Federal Circuit. It argues that current institutional critiques of the Federal Circuit can be best framed as two concerns: (1) the relationship of the Federal Circuit with lower courts and agencies, and (2) the adjudicative “form” of patent law.

39. Approximately 55% of the Federal Circuit’s docket consists of patent cases. See Tony Dutra, *‘Introspective Look’ at Federal Circuit Highlights Breadth of Court’s Docket*, 77 PAT. TRADEMARK & COPYRIGHT J. 560, 560 (2009).

40. See, e.g., Rochelle Cooper Dreyfuss, *The Federal Circuit: A Case Study in Specialized Courts*, 64 N.Y.U. L. REV. 1, 1 (1989) (proposing that the Federal Circuit “shift its focus from patent law to competition law more generally” to be more efficient).

41. See, e.g., Rai, *Engaging Facts and Policy*, *supra* note 8, at 1123; Thomas, *supra* note 7, at 796; Wood, *supra* note 6, at 1.

A. RELATIONSHIP WITH LOWER COURTS AND AGENCIES

Several commentators have criticized the Federal Circuit for its unusual, non-deferential relationship with both district courts and the two agencies (PTO and ITC) whose decisions it reviews. The critique can be summed up as follows: the Federal Circuit shows little deference to district courts on questions of fact—despite Rule 52(a) of the Federal Rules of Civil Procedure, which requires a more deferential clearly erroneous standard of review of factual issues—and shows little deference to the Patent and Trademark Office on technical issues—displaying considerable reluctance to applying the Administrative Procedure Act (“APA”).⁴² The next two sections analyze these critiques in depth.

1. *Relationship with Lower Courts*

Academic commentators have criticized the Federal Circuit for showing trial courts less deference than is required by traditional doctrines of appellate review. Under Rule 52(a),⁴³ an appeals court reviews legal conclusions *de novo*, but reviews factual findings under the more deferential *clearly erroneous* standard.⁴⁴ The Federal Circuit, however, has sidestepped this division of labor between trial and appellate courts, either by interpreting questions of fact, or mixed questions of law and fact, as purely questions of law, or by applying a *de novo* standard of review to inquiries the Federal Circuit itself has characterized as factual.⁴⁵ Commentators have blamed this behavior for creating increased unpredictability and uncertainty in patent claim construction⁴⁶ and, more broadly, for divorcing patent law from the context-specific needs of different innovation communities.⁴⁷

42. See, e.g., Rai, *Engaging Facts and Policy*, *supra* note 8; Stuart Minor Benjamin & Arti K. Rai, *Who's Afraid of the APA?, What the Patent System Can Learn from Administrative Law*, 95 GEO. L.J. 269 (2007); Craig Allen Nard, *Deference, Defiance, and the Useful Arts*, 56 OHIO ST. L.J. 1415, 1421–22 (1995) (“[T]he court has adopted a highly questionable, if not erroneous, *de novo* standard of review [of PTO denials of patentability decisions].”).

43. FED. R. CIV. P. 52(a).

44. *Id.* In one of its only two patent opinions of the 1980s, the Supreme Court required the Federal Circuit to adhere to the same standard of review of factual questions as other Article III courts. See *Dennison Mfg. Co. v. Panduit Corp.*, 475 U.S. 809 (1986) (*per curiam*).

45. See, e.g., Rai, *Engaging Facts and Policy*, *supra* note 8, at 1056 (noting that the Federal Circuit “has subjected trial court fact finding in infringement or declaratory judgment actions to a level of review that is contrary to traditional principles of appellate review”); William F. Lee & Anita K. Krug, *Still Adjusting to Markman: A Prescription for the Timing of Claim Construction Hearings*, 13 HARV. J.L. & TECH. 55, 67 (1999).

46. See, e.g., Jonas Anderson & Peter Menell, *Informal Deference: A Historical, Empirical, and Normative Analysis of Patent Claim Construction*, 108 NW. U. L. REV. 1, 1 (2014) (“[T]he Federal Circuit’s adherence to the *de novo* standard has frustrated district courts’ distinctive capabilities for apprehending and resolving the factual disputes inherent in claim

The Federal Circuit's proclivity for interpreting arguably factual questions as issues of law is exemplified in claim construction—the determination of the “metes and bounds” of the inventive territory. Claim construction differs from statutory interpretation—an issue of law—in that patent claims are analyzed from the perspective of a “person having ordinary skill in the art” (“PHOSITA”).⁴⁸ Thus, there is a good argument that expert testimony regarding how a PHOSITA would understand a claim should play an important role in ensuring agreement between the relevant community of innovators and the ultimate interpretation of the claim language at issue. Evaluating expert evidence and its credibility is a task traditionally considered “fact-finding,” and therefore the province of the trial court.⁴⁹ Nevertheless, the Federal Circuit has not interpreted any aspect of claim construction—including the use and evaluation of expert testimony—as fact-finding.⁵⁰

This interpretive stance has generated vigorous criticism from both academics and district court judges. One scholar observed that the Federal Circuit is simply ignoring both the “significant role for facts in claim construction”⁵¹ as well as crucial Supreme Court language in *Markman v. Westview Instruments, Inc.* that characterizes claim construction as a “mongrel practice” involving both legal and factual issues.⁵² District court judges have also been critical of the Federal Circuit's penchant for de novo review in

construction determinations . . .”); David L. Schwartz, *Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases*, 107 MICH. L. REV. 223, 259 (2008) (noting that the high reversal rates in claim construction decrease incentives for trial judges to learn, increases the unpredictability of appellate outcomes, and decreases the possibility of settlement); Christian A. Chu, *Empirical Analysis of the Federal Circuit's Claim Construction Trends*, 16 BERKELEY TECH. L.J. 1075, 1113 (2001) (noting that “claim construction review plays a significant role in the likelihood of reversal”).

47. See, e.g., Rai, *Engaging Facts and Policy*, *supra* note 8, at 1106–07; Thomas, *supra* note 7, at 796; Craig Allen Nard & John F. Duffy, *Rethinking Patent Law's Uniformity Principle*, 101 NW. U. L. REV. 1619, 1623–25 (2007).

48. 35 U.S.C. § 103(a) (2006); Dreyfuss, *supra* note 10, at 29–32.

49. See, e.g., Moberly *ex rel.* Moberly v. Sec'y of Health & Human Servs., 592 F.3d 1315, 1326 (Fed. Cir. 2010) (“Finders of fact are entitled—indeed, expected—to make determinations as to the reliability of the evidence presented to them and . . . as to the credibility of the persons presenting that evidence.”).

50. *Cybor Corp. v. FAS Techs., Inc.* 138 F.3d 1448, 1451 (Fed. Cir. 1998) (holding that, based on the Supreme Court's holding in *Markman*, “claim construction, as a purely legal issue, is subject to de novo review on appeal”).

51. Rai, *Engaging Facts and Policy*, *supra* note 8, at 1048.

52. *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 378 (1996). Anderson and Menell have expressed similar concerns and urged the court to adopt a “hybrid standard” of review that “would defer to trial judges' factual determinations” on how a PHOSITA would understand technical terms but “would retain de novo authority over whether the trial court's factual finding inappropriately overrides more specific intrinsic indications of the patent's scope.” Anderson & Menell, *supra* note 46, at 1–2.

claim construction. For example, Chief Judge Saris of the District Court of Massachusetts argued that “[t]here should be more deference [to the district judge on claim construction] particularly when the district judge takes expert testimony or receives other extrinsic evidence.”⁵³ And while serving at the District Court for the Northern District of Ohio, Judge O’Malley (now a judge at the Federal Circuit) observed: “it is a hard pill to swallow as a district judge that, after seeing the experts, and hearing the experts, our efforts to answer those questions are subject to a completely de novo review and a blank record.”⁵⁴ Indeed, on January 20, 2015 the Supreme Court reversed the Federal Circuit on this particular issue, holding in *Teva Pharmaceuticals U.S.A., Inc. v. Sandoz, Inc.* that Rule 52(a)(6) requires the Federal Circuit to review subsidiary factual matters made in the course of claim construction of a patent claim for “clear error.”⁵⁵

The Federal Circuit’s tendency to review de novo arguably factual issues extends beyond claim construction. For example, two cases heard by the Supreme Court in its 2013 term involved the proper standard of review of district court decisions to award attorneys’ fees in exceptional cases pursuant to 35 U.S.C. § 285.⁵⁶ In *Highmark Inc. v. Allcare Health Management Systems, Inc.*, petitioners argued that the Federal Circuit had “arrogate[d] th[e] responsibility” to “apply the fact-dependent legal standard,” and thus “improperly divid[ed] labor between the trial courts and courts of appeal.”⁵⁷ In *Octane Fitness, LLC v. Icon Health & Fitness, Inc.*, petitioners similarly argued that the Federal Circuit had “improperly appropriate[d] a district court’s discretionary authority to award attorney fees to prevailing accused infringers in contravention of statutory intent and this Court’s precedent.”⁵⁸ In both cases, the Supreme Court sided with petitioners, striking down the Federal Circuit’s de novo standard of review.⁵⁹ The Court emphasized that the Federal Circuit had relied on an “unduly rigid” framework that

53. See, e.g., Kathleen M. O’Malley, Patti Saris & Ronald H. Whyte, *A Panel Discussion: Claim Construction from the Perspective of the District Judge*, 54 CASE W. RES. L. REV. 671, 679–80 (2004) (statement of Hon. Patti Saris).

54. *Id.* at 680 (statement of Hon. Kathleen O’Malley); Nard & Duffy, *supra* note 47, at 1620.

55. *Teva Pharma. U.S.A., Inc. v. Sandoz, Inc.*, 135 S. Ct. 831 (2015).

56. *Highmark Inc. v. Allcare Health Mgmt. Sys., Inc.*, 134 S. Ct. 1744 (2014); *Octane Fitness, LLC v. ICON Health & Fitness, Inc.*, 134 S. Ct. 1749 (2014).

57. Petition for a Writ of Certiorari at 4, *Highmark*, 134 S. Ct. 1744 (2014) (No. 12-1163), 2013 WL 1209137.

58. Petition for a Writ of Certiorari at i, *Octane*, 134 S. Ct. 1749 (2014) (No. 12-1184), 2013 WL 1309080.

59. *Highmark*, 134 S. Ct. at 1748–49.

“impermissibly encumbers the statutory grant of discretion to district courts.”⁶⁰

Commentators have also criticized the Federal Circuit for acting as a fact-finder itself—even for issues that it recognizes as plainly factual matters. For example, even though the Federal Circuit considers the ultimate question of patent infringement to be a factual determination,⁶¹ it often issues a ruling on infringement following claim construction, rather than remanding the case to the district court for a new trial.⁶² In this same context, Arti Rai has criticized the court’s penchant for “simply declar[ing] that there can be no factual dispute with respect to infringement.”⁶³ And in decisions concerning patent validity, including non-obviousness and disclosure determinations, Rai contends that the Federal Circuit has “merely paid lip service to deference,” while actually substituting its own fact-finding for that of the district court.⁶⁴

Federal Circuit judges themselves are not all of a piece: some have opposed a purely de novo standard of review in claim construction and other arguably factual inquiries, such as whether a case is exceptional.⁶⁵ Yet, the relatively few dissents in claim construction and the recent Federal Circuit decision in *Lighting Ballast* confirming the continued validity of a purely de novo standard of review, imply a continued tendency to review arguably factual issues de novo.⁶⁶ Indeed, it is telling that Federal Circuit judges have characterized their own Court’s behavior as a “temptation to label everything legal and usurp the province of the fact finder with our manufactured de novo review.”⁶⁷

60. *Octane Fitness*, 134 S. Ct. at 1755.

61. *Hilton Davis Chem. Co. v. Warner-Jenkinson Co.*, 62 F.3d 1512, 1520–21 (Fed. Cir. 1995) (en banc) (“infringement, whether literal or under the doctrine of equivalents, is a question of fact”).

62. See, e.g., William C. Rooklidge & Matthew F. Weil, *Judicial Hyperactivity: The Federal Circuit’s Discomfort with Its Appellate Role*, 15 BERKELEY TECH. L.J. 725, 739–40 (2000).

63. Rai, *Engaging Facts and Policy*, *supra* note 8, at 1060.

64. *Id.* at 1063.

65. For example, Judge Mayer, joined by Judge Newman, wrote a dissent in *Phillips v. AWH Corp.*, 415 F.3d 1303, 1330 (Fed. Cir. 2005) (Mayer, J., dissenting) reiterating his continued frustration with “the futility, indeed the absurdity, of this court’s persistence in adhering to the falsehood that claim construction is a matter of law devoid of any factual component.” And in *Highmark, Inc. v. Allcare Health Mgmt. Sys.*, 701 F.3d 1351, 1357 (Fed. Cir. 2012), five judges (Chief Judge Rader and Judges Moore, O’Malley, Reyna, and Wallach, dissenting from the Federal Circuit’s denial of rehearing en banc) argued that the court’s de novo standard of review in exceptional cases “invades the province of the fact finder.”

66. *Lighting Ballast Control v. Philips Elecs. N. Am. Corp.*, 744 F.3d 1272, 1276–77 (Fed. Cir. 2014) (en banc) (emphasis added).

67. *Highmark*, 701 F.3d at 1362 (Fed. Cir. 2012).

Whether a more deferential standard of review is normatively desirable is an open question, given the Federal Circuit's mandate to maintain uniformity, as well as its unique knowledge of patent law. Indeed, Rochelle Dreyfuss has suggested that the "Federal Circuit's unique responsibility towards patent law argues for a broader scope of review over fact finding."⁶⁸ Nevertheless, it is clear that the Federal Circuit has chosen not to defer to the district court on issues for which there is a strong case for deference under traditional principles of appellate review.

2. *Relationship with the Patent and Trademark Office and the International Trade Commission*

The Federal Circuit is also an outlier in its review of agency action: it has adopted a less deferential standard of review of administrative fact-finding and statutory interpretation than all other Article III appellate courts. The Federal Circuit interacts routinely with two agencies that are thought to possess a degree of expertise in patent law: the Patent and Trademark Office ("PTO") and the International Trade Commission ("ITC"). The Federal Circuit has an asymmetric relationship with the PTO: it can review directly its *denials* of patent protection, but not its patent *grants*. The latter only reach the Federal Circuit through an appeal from a district court decision.⁶⁹ The ITC makes trade-focused patentability decisions under section 337 of the Tariff Act, which gives the ITC authority to grant broad exclusion orders to entities whose patents have been infringed by imported goods.⁷⁰ ITC Section 337 determinations may be appealed directly to the Federal Circuit.⁷¹

In reviewing PTO patent denials, the Federal Circuit has not followed the traditional deference structure that appellate courts employ with administrative agencies. Under section 706 of the Administrative Procedure Act ("APA"), courts review administrative fact-finding under the highly deferential "arbitrary or capricious" or "unsupported by substantial evidence" standards.⁷² Yet, until the Supreme Court intervened in *Dickinson v. Zurko*,⁷³ the Federal Circuit maintained that the APA did not apply to its

68. Dreyfuss, *supra* note 40, at 61–62. In her more recent writings, however, Dreyfuss emphasizes that an examination of the history of the Federal Circuit underscores the emergence of "a problem that was largely unforeseen . . . the high cost of eliminating intercourt debate from the adjudicatory system." Dreyfuss, *supra* note 12, at 507. Dreyfuss favors increased deference to the PTO and specialization at the trial level as a way to increase institutional dialogue. *Id.* at 532–36.

69. 28 U.S.C. § 1295(a)(4) (2006).

70. 19 U.S.C. § 1337(a) (2006).

71. 19 U.S.C. § 1337(c) (2006).

72. 5 U.S.C. § 706(2)(A), (E) (2006).

73. *Dickinson v. Zurko*, 527 U.S. 150 (1999).

reviews of the PTO's findings of fact, choosing instead to apply a more rigorous standard of review.⁷⁴ And although the Supreme Court's *Zurko* decision held that the APA did apply to the Patent Act, commentators have observed that the Federal Circuit has continued to display considerable resistance to applying APA standards of review to its patent docket.⁷⁵

The Federal Circuit has repeatedly refused to grant deference to the PTO's substantive interpretations of the Patent Act.⁷⁶ Rather, it considers the PTO to have only procedural—not substantive—rule-making authority with respect to the Patent Act.⁷⁷ The Federal Circuit has also retained the power of adjudicating *de novo* whether a particular rule is procedural or substantive. As a consequence, the Federal Circuit reviews the PTO's findings of (substantive) law *de novo*, and its findings of fact under a clearly erroneous standard.⁷⁸ Many academic commentators have criticized this division of labor, arguing for greater deference to the PTO—for example, by deferring to the PTO's determinations of whether a rule is substantive or procedural,⁷⁹ or by granting the PTO substantive rule making authority and *Chevron* deference to its adjudicative decisions.⁸⁰

As elaborated further in Section III.A, the relationship between the Federal Circuit and the PTO can be described and understood as one of *jurisdictional competition* between two expert communities for control over patent law. In fact, this type of competition is expected under a sociological model of institutionalized communities of experts.

74. *In re Zurko*, 142 F.3d 1447, 1459 (Fed. Cir. 1998).

75. *See, e.g.*, Rai, *Engaging Facts and Policy*, *supra* note 8, at 1055.

76. *See, e.g.*, *Tafas v. Doll*, 559 F.3d 1345, 1352 (Fed. Cir. 2009) (holding that the PTO does not have “any general substantive rulemaking power”); *Brand v. Miller*, 487 F.3d 862, 869 n.3 (Fed. Cir. 2007) (holding that the PTO “does not earn *Chevron* deference on questions of substantive patent law”).

77. *Merck & Co. v. Kessler*, 80 F.3d 1543, 1549–50 (Fed. Cir. 1996) (noting that the PTO lacks the ability to promulgate rules on the core patentability standards that carry the force of law).

78. *In re Lueders*, 111 F.3d 1569, 1571–72 (Fed. Cir. 1997) (“We review the Board’s ultimate legal conclusion . . . *de novo* . . . [and] underlying findings of fact for clear error.”); *In re Zurko*, 142 F.3d 1447, 1449 (Fed. Cir. 1998) (“We believe section 559 of the [APA] permits . . . our continued application of the clearly erroneous standard in our review of these fact-findings.”).

79. *See, e.g.*, Joseph Scott Miller, *Substance, Procedure and the Divided Patent Power*, 63 ADMIN. L. REV. 31, 49–50 (2011).

80. *See, e.g.*, Michael J. Burstein, *Rules for Patents*, 52 WM. & MARY L. REV. 1747, 1751 (2011); Jonathan S. Masur, *Regulating Patents*, in 2010 THE SUPREME COURT REVIEW 275, 279 (Dennis Hutchinson et al. eds., 2011); Melissa F. Wasserman, *The Changing Guard of Patent Law: Chevron Deference for the PTO*, 54 WM. & MARY L. REV. 1959, 1998 (2013).

The Federal Circuit has similarly refused to grant either *Chevron* or *Skidmore* deference to patent decisions from the ITC.⁸¹ Doctrinally, the case for deference to the ITC differs from the case for deference to the PTO. Since the ITC has interpretative authority over the Tariff Act only, deference would only be warranted under the APA and *Chevron* if the ITC is interpreting the Tariff Act—but not the Patent Act—when making patentability determinations.⁸² Normatively, commentators are divided on whether deference to ITC is desirable. For example, focusing on the importance of avoiding a fragmented patent regime, John Thomas argues against ITC deference in patentability and infringement determinations.⁸³ In contrast, Sapna Kumar has argued that considerations of institutional competence and political accountability favor granting *Chevron* deference to ITC patentability and infringement decisions.⁸⁴

Leaving aside whether increased deference is normatively desirable, what remains clear is that the Federal Circuit is an outlier among Article III courts in its review of agency action.⁸⁵ Significantly, the Federal Circuit has only attempted to arrogate power over fact-finding and statutory interpretation on patent law issues, while routinely granting APA and *Chevron* deference to agencies that do not handle patent disputes, and to PTO and ITC on non-patent matters.⁸⁶

81. See, e.g., *Linear Tech. Corp. v. Int'l Trade Comm'n*, 566 F.3d 1049, 1068 (Fed. Cir. 2009); *Honeywell Int'l, Inc. v. Int'l Trade Comm'n*, 341 F.3d 1332, 1338 (Fed. Cir. 2003); *Mead Corp. v. United States (Mead II)*, 185 F.3d 1304, 1305 (Fed. Cir. 1999); *Mead Corp. v. United States (Mead III)*, 283 F.3d 1342, 1344 (Fed. Cir. 2002). See also Kumar, *supra* note 9, at 1547; Kali Murray, *First Things First: A Principled Approach to Patent Administrative Law*, 42 J. MARSHALL L. REV. 29, 48–54 (2008).

82. See Kumar, *supra* note 9, at 1562–63 (arguing that the ITC is interpreting the Tariff Act when making patentability determinations to decide whether to grant an exclusion order); see also *Process Patents: Hearing Before the S. Comm. on the Judiciary, 110th Cong.* 86–87 (2007) (statement of John R. Thomas, Professor of Law, Georgetown University) (testifying that the ITC “is not charged with interpreting the Patent Act” whenever it makes patent-related determinations).

83. *Process Patents: Hearing Before the S. Comm. on the Judiciary, 110th Cong.* 6, 87 (2007) (statement of John R. Thomas, Professor of Law, Georgetown University).

84. See Kumar, *supra* note 9, at 1587, 1592.

85. See, e.g., Banks Miller & Brett Curry, *Experts Judging Experts: The Role of Expertise in Reviewing Agency Decision Making*, 38 LAW & SOC. INQUIRY 55, 55–56 (2013) (showing that courts are generally deferential to the decision-making of federal agencies).

86. Kumar, *supra* note 9, at 1566 (noting that the Federal Circuit grants both APA and *Chevron* deference “when patents are not at issue”); Craig Allen Nard, *Deference, Defiance and the Useful Arts*, 56 OHIO ST. L.J. 1415, 1430 (1995) (“The impact of *Chevron* has been lost on the Federal Circuit as it relates to the BPAP’s patentability determinations; whereas just the opposite can be said about the Federal Circuit’s nonpatent administrative caseload.”).

B. ADJUDICATIVE “FORM” OF PATENT LAW

In a seminal article, Duncan Kennedy distinguished two “different rhetorical modes” of private law adjudication regarding the *form* of legal decisions.⁸⁷ The first rhetorical mode “favors the use of clearly defined, highly administrable, general rules,” while “the other supports the use of equitable standards producing ad hoc decisions with relatively little precedential value.”⁸⁸ By all accounts, the Federal Circuit has consistently favored the use of clear and inflexible general rules.⁸⁹ Time and again, the court has attempted to distill patentability inquiries into highly administrable rules that eschew contextual analysis and that limit the ability of lower courts to adjust their decisions to the circumstances of the case. The Federal Circuit’s systematic preference for formal, rigid rules over flexible standards, many commentators believe, has contributed to patent law’s disconnect from the needs of communities of innovators working in a quickly-evolving technological environment.

One of the most prominent, and most criticized, examples of Federal Circuit rule formalism is the Court’s development of the “teaching, suggestion, or motivation” (“TSM”)⁹⁰ test for determining whether an invention is non-obvious under section 103 of the Patent Act.⁹¹ When an invention involves combining two or more references, the TSM test required that at least one of those references contain information that would suggest, teach, or motivate a PHOSITA to combine the references at issue.⁹² Although designed to avoid hindsight bias, the TSM test ultimately prevented consideration of contextual factors, such as tacit knowledge in the relevant

87. Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685, 1685 (1976).

88. *Id.*

89. See, e.g., Dreyfuss *supra* note 12, at 530; Holbrook, *supra* note 7, at 123 (“In recent years, the . . . Federal Circuit has embraced the use of bright-line, formalistic rules . . .”); Nard & Duffy, *supra* note 47, at 1644 (suggesting that the Federal circuit “has retreated into its own legal formalisms at the expense of gaining a good understanding of industrial and technological needs”); Rai, *Engaging Facts and Policy*, *supra* note 8, at 1040 (“[T]he Federal Circuit has substituted formalist decisionmaking for the fact-specific, policy-oriented analysis that is required by the open-ended language of the patent statute.”); Thomas, *supra* note 7, at 774 (“[T]he Federal Circuit has embraced an increasingly formal jurisprudence.”); see also Lee, *supra* note 7, at 21–22 (arguing that the Federal Circuit resorts to inflexible rules to avoid engaging with complex technologies).

90. KSR Int’l Co. v. Teleflex Inc., 550 U.S. 398, 399 (2007).

91. 35 U.S.C. § 103(a) (2006). See, e.g., Laura G. Pedraza-Fariña, *Patent Law and the Sociology of Innovation*, 2013 WIS. L. REV. 813, 823–26 (2013) (describing the evolution of the Federal Circuit’s formalist obviousness jurisprudence and contrasting it with the contextual approach embraced by the Supreme Court).

92. Pedraza-Fariña, *supra* note 91, at 825.

scientific community, that would lead one of ordinary skill in the art to combine references even absent an explicit indication to do so.⁹³ This is essentially the conclusion reached by the Supreme Court in *KSR International Co. v. Teleflex, Inc.* when it rejected the TSM test as the sole obviousness analysis, and replaced it with a case-by-case focus on what scientists and inventors would know or could develop during routine research.⁹⁴ Interestingly, although one might expect the Federal Circuit in the wake of *KSR* to begin placing much more attention on defining the attributes of a person having ordinary skill in the art,⁹⁵ the court still rarely does so.⁹⁶ Rather, the court appears to be sliding back into pre-*KSR* rule formalism.⁹⁷

While the court's obviousness jurisprudence is one of the most salient examples of its reliance on rigid rules, it is by no means the only one. The Federal Circuit has favored bright-line rules over flexible standards in determining whether an invention is novel, in patentable subject matter determinations, in establishing remedies and, more recently, in its extraordinary case jurisprudence.⁹⁸ In a key patentable subject matter decision, *In re Bilski*, the Federal Circuit developed the "machine or transformation test" as the "sole test" of patentability for process claims.⁹⁹ Under this test, a process may constitute patentable subject matter if it (1)

93. *Id.* at 823–25; Rebecca S. Eisenberg, *Obvious to Whom? Evaluating Inventions from the Perspective of PHOSITA*, 19 BERKELEY TECH. L.J. 885, 896 (2004) (arguing that the TSM test had the "practical effect of excluding from the analysis the tacit knowledge ordinary practitioners commonly possess").

94. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 421 (2007) ("A person of ordinary skill is also a person of ordinary creativity, not an automaton."); see also Daralyn J. Durie & Mark A. Lemley, *A Realistic Approach to the Obviousness of Inventions*, 50 WM. & MARY L. REV. 989, 999 (2008) ("[T]he one consistent strand that runs through the opinion is a rejection of rigid rules, replaced with a case-by-case focus on what actual scientists in the field would know or could develop with ordinary inventive skill.").

95. See Durie & Lemley, *supra* note 94, at 1000–03 (arguing that, in the aftermath of *KSR*, courts should pay increased attention to "the way in which PHOSITAs work in the real world").

96. See, e.g., Emer Simic, *The TSM Test is Dead! Long Live the TSM Test! The Aftermath of KSR, What Was All the Fuss about?* 37 AIPLA Q.J. 227, 247–48 (2009) (discussing the Federal Circuit's unwillingness to clearly define a PHOSITA in multiple cases).

97. See, e.g., Pedraza-Fariña, *supra* note 91, at 862–63 (showing how "in its more recent opinions, the Federal Circuit has taken a narrower and more formalistic view of analogous art," which is a key step in the obviousness inquiry).

98. See Thomas, *supra* note 7, at 778–81 (novelty); *In re Bilski*, 545 F.3d 943, 955–56 (Fed. Cir. 2008) (patentable subject matter); *Apple, Inc. v. Samsung Elec. Co.*, 678 F.3d 1314, 1335 (Fed. Cir. 2012) (remedies); *Icon Health & Fitness, Inc. v. Octane Fitness, LLC*, 496 Fed. App'x 57, 65 (Fed. Cir. 2012) (exceptional case jurisprudence).

99. *In re Bilski*, 545 F.3d 943, 955–56 (Fed. Cir. 2008). The "machine-or-transformation" test was first articulated in *Gottschalk v. Benson*, 409 U.S. 63, 69–70 (1972).

utilizes a particular machine or apparatus or (2) transforms an object into a different state or thing.¹⁰⁰ On appeal, however, the Supreme Court refused to limit the patentability inquiry to the machine or transformation test, calling it a “categorical rule” that would “frustrate the purposes of patent law” by hindering the adoption of technological advances.¹⁰¹ Similarly, prior to the Supreme Court’s decision in *eBay Inc. v. MercExchange, LLC*, the Federal Circuit relied on a “general rule” that injunctions should issue in patent infringement cases absent exceptional circumstances.¹⁰² The Supreme Court struck down this general rule and held that patent infringement cases were not an exception to “the traditional four factor test applied by courts of equity when considering whether to award permanent injunctive relief to a prevailing plaintiff.”¹⁰³ Most recently, the Supreme Court rejected as “rigid and mechanical” another bright-line rule that the Federal Circuit had developed in its exceptional case jurisprudence.¹⁰⁴

Commentators are split regarding the normative desirability of rule-formalism. Practicing patent attorneys have by and large welcomed the Federal Circuit’s turn to rules as increasing predictability,¹⁰⁵ while a majority of academic commentators have denounced it as inconsistent with patent law’s goal of promoting innovation.¹⁰⁶ As Part III elaborates, predictions from the model of expert decision-making described in this article undercut the assumption that rule formalism will generate uniform decisions. Specifically, communities of experts are expected to look for ways to free themselves from the very rules they create to constrain their subordinates when these rules do not accord with their intuitions, thus making a rule-

100. *In re Bilski*, 545 F.3d 943, 955–56 (Fed. Cir. 2008).

101. *Bilski v. Kappos*, 561 U.S. 593, 605 (2010).

102. *eBay Inc. v. MercExchange, LLC*, 547 U.S. 388, 388 (2006) (disapproving of the use of a “general rule”); *Apple*, 678 F.3d at 1335 (noting that the Supreme Court rejected the “general rule” in *eBay*).

103. *eBay*, 547 U.S. at 388.

104. *Octane Fitness, LLC v. ICON Health & Fitness, Inc.*, 134 S. Ct. 1749, 1755 (2014). *See also* Dreyfuss *supra* note 12, at 512–15 (describing “the debate between the Supreme Court and the Federal Circuit over the use of bright line rules” and the Supreme Court’s rejection of the Federal Circuit’s rule formalism).

105. *See, e.g.*, Russell B. Hill & Frank P. Cote, *Ending the Federal Circuit Crapshoot: Emphasizing Plain Meaning in Patent Claim Interpretation*, 42 IDEA 1, 2 (2002) (extolling “the need for uniform application of the presumption in favor of plain meaning in patent claim interpretation”); Lee & Evans, *supra* note 8, at 7 (discussing the historical right to a patent jury trials as “reaffirmed” in *Markman*); Victoria Slind-Flor, *Federal Circuit Judged Flawed*, NAT’L L.J., Aug. 3, 1998 (noting that some have accused the court of “unpredictability”); Thomas *supra* note 7, at 794 (noting that patent lawyers were a powerful lobby that advocated for clear rules in patent law).

106. *See* references listed *supra* note 7.

based system much less predictable than would otherwise be anticipated. That the Federal Circuit does indeed routinely break its own rules helps explain, at least in part, the puzzling observation that the Federal Circuit has in fact failed to bring uniformity and predictability to its docket.¹⁰⁷

* * *

This Part has provided a synthesis of two lines of critique of the Federal Circuit as an institution. First, the Federal Circuit consistently grants less deference to decisions by trial courts and agencies than other Article III appellate courts typically grant, and, arguably, less deference than is required by Rule 52(a) and the APA, respectively. Second, the Federal Circuit prefers inflexible rules of decision to flexible standards, even though the rapid advances in technology that characterize the field of patent law may call for a more contextual, case-by-case approach. The next Part introduces the reader to studies in the sociology of expertise and adapts insights from this literature to the hierarchical court system.

III. APPROACHES TO EXPERT DECISION-MAKING

This Part synthesizes and brings together two approaches to the study of expertise generally: the institutional approach and the embodied approach.

A. THE INSTITUTIONAL APPROACH TO EXPERTISE: SOCIOLOGY OF THE PROFESSIONS

An important approach to the study of expertise focuses on how expertise is organized and controlled in society. This line of research studies the development of professional organizations and other institutionalized forms of expertise, as well as how expert institutions interact with one another.¹⁰⁸ From this body of work emerge two key insights with important consequences for the study of expert courts. First, groups of experts seek to maximize their control and autonomy over both the development of a system of expert abstract knowledge and its range of application. These expert groups, however, exist in an ecosystem of other expert groups who apply their distinct knowledge systems to the same sets of problems. For example, both physicians and psychologists compete over the application of

107. *See infra* Subsection IV.B.2.

108. *See generally* ANDREW ABBOTT, *THE SYSTEM OF PROFESSIONS: AN ESSAY ON THE DIVISION OF EXPERT LABOR* (1988); ELIOT FRIEDSON, *PROFESSIONALISM: THE THIRD LOGIC* (2001); MAGALI SARFATTI LARSON, *THE RISE OF PROFESSIONALISM: A SOCIOLOGICAL ANALYSIS* (1977); Elizabeth H. Gorman & Rebecca L. Sandefur, "Golden Age," *Quiescence, and Revival: How the Sociology of Professions Became the Study of Knowledge-Based Work*, 38 *WORK & OCCUPATIONS* 275 (2011).

their expert knowledge to mental health issues. The overlap between communities leads to competition for jurisdictional autonomy and control. Second, professions employ two key mechanisms in their competition for jurisdictional control: (1) *framing* of overlapping problems as being best solved through a particular expert community's abstract knowledge system, and (2) *codification* of a portion of that abstract knowledge into rules of easy application. I explore these two themes below.

Through a series of studies of professional organizations, such as those of doctors and psychologists, Andrew Abbott, Eliot Friedson, and others have theorized that the essence of a profession is to seek maximal autonomy and control over the set of abstract principles within its "jurisdiction."¹⁰⁹ In turn, a profession's jurisdiction is simply those tasks the profession considers to be (and that it convinces society should be) within its body of expert knowledge.¹¹⁰ Take, for example, the field of childhood health and development. Doctors have jurisdiction over the diagnosis and treatment of physical ailments, but a number of other professional groups have increasingly claimed jurisdiction over a subset of childhood physical problems, such as physical therapists, occupational therapists, speech language pathologists, and developmental therapists. Because of their overlapping claims to jurisdiction, these groups will often compete with each other and with physicians over the appropriate diagnosis and treatment of a problem. Having complete jurisdictional control means having the power to define and classify a problem, to define and apply the correct treatment, and to evaluate the treatment's success.¹¹¹ These studies define the term "profession" quite broadly, to encompass any exclusive or semi-exclusive community of experts that develops abstract knowledge and applies it to particular cases.¹¹²

In essence, the claim is straightforward: because organized groups of experts will seek to maintain control over their body of knowledge (composed of abstract principles), they will reject claims by those outside the profession to legitimately dictate what those professionals do or how they do it. Yet, this claim has crucial implications for understanding the interaction among expert communities and between experts and non-experts. This initial

109. Abbott, *supra* note 108, at 71 ("Full jurisdictional claims are, in general, the goal of all other types of settlements."); ELIOT FREIDSON, PROFESSIONAL POWERS: A STUDY OF THE INSTITUTIONALIZATION OF FORMAL KNOWLEDGE (1986).

110. Abbott, *supra* note 108, at 59.

111. *Id.*

112. *See, e.g.*, Gorman & Sandefur, *supra* note 108, at 277 ("In the eyes of contemporary scholars, the commonalities between traditional professions and new forms of knowledge-based work are more important than the differences.").

claim immediately implies another: professions do not exist in isolation, but are embedded in an ecosystem where they compete with one another for jurisdictional control.¹¹³ In turn, dissecting the mechanisms by which such competition takes place is important for understanding expert community dynamics. The following subsections explore two such mechanisms: (1) competing framings of tasks and problems as best solved within a profession's system of abstract knowledge, and (2) codification of (at least a portion of) such abstract knowledge.

1. *Competing Framings of Tasks and Problems*

Professions seek to gain and maintain jurisdictional control through the development and control of a *system of expert abstract knowledge*, which only members of the profession have access to and can apply to specific cases.¹¹⁴ For example, different medical specialties have developed abstract knowledge systems that correlate symptoms with disease diagnoses, mechanistic explanations for the disease, and appropriate treatments.¹¹⁵ Law is itself built on different systems of abstract knowledge. In patent law, concepts such as “a person of ordinary skill in the art,”¹¹⁶ “conception,”¹¹⁷ and “non-obviousness,”¹¹⁸ to name only a few, are elements in an abstract knowledge system designed to ultimately incentivize innovation.

Professions use abstract knowledge to classify and offer solutions for tasks and problems. Tasks can be conceptualized as having both objective and subjective elements. Objective elements are features of a problem so broadly agreed upon as constitutive of that problem that they come to represent fixed characteristics not easily reinterpreted.¹¹⁹ The most obvious objective element of a problem is given by its natural or factual characteristics. For example, any treatment for alcoholism is bounded by objective characteristics of alcohol consumption itself, such as loss of fine motor skills, as well as coarse motor skills and sensory function at high consumption levels.¹²⁰ Thus, objective elements of tasks require that a

113. See, e.g., Abbott, *supra* note 108, at 115 (discussing how external forces impact the development of expert communities).

114. See, e.g., *id.* at 70.

115. See, e.g., THE MERCK MANUAL OF DIAGNOSIS AND THERAPY (Robert S. Porter et al. eds., 19th ed. 2011) (“As it has for over 110 years *The Manual* provides health care practitioners . . . with straightforward, practical explanations of ‘what to do’ to diagnose and treat . . . conditions.”)

116. 35 U.S.C. § 103 (2006).

117. *Id.* § 201(g).

118. *Id.* § 103.

119. See, e.g., Abbott *supra* note 108, at 35–40.

120. Abbott, *supra* note 108, at 37–38.

profession's definition of a problem remain closely linked to that problem's fixed attributes. Subjective elements, on the other hand, are *framings* of a particular problem claimed by a particular profession.

Professions compete for jurisdictional control by framing problems as best solved within their unique abstract knowledge system. But there are often multiple ways to frame a problem's subjective qualities. Thus, the act of classifying a problem creates the arena where jurisdictional struggles take place. A classic example of jurisdictional competition through the control of a system of abstract knowledge concerns the struggle among clergy, medicine, psychiatry, and criminal law to define and treat alcoholism. Each one of these four communities sought to frame alcoholism—and thus to control the market for treatment—according to their own abstract knowledge systems:

At first [alcoholism was] a moral and spiritual problem; ministers were the relevant experts. The doctors soon attacked, substituting the claim of cure for the clergyman's mere condemnation and forgiveness. In the late nineteenth century, the problem was pronounced a legal one, although the lawyers and the police dealt with alcoholism simply by incarcerating it. The psychiatrists also claimed alcoholism in this period.¹²¹

Alcoholism framed as a mental disorder concerning addiction and impulse control grants primary jurisdictional control over treatment to psychologists or psychiatrists; framed as a problem involving neurotransmitter hypersensitivity, it grants primary jurisdictional control to physicians.

2. *Rule-Making or Codification*

A second, and complementary, form of control involves the *codification of abstract knowledge*. Codification, or rule making, allows professions to delegate work to subordinate professions while maintaining control over the abstract principles that create those rules.¹²² Codification allows expert communities to expand their jurisdiction by enlisting other—subordinate—communities

121. *Id.* at 37. A more recent iteration of this type of competition is between “scientific psychiatry,” embodied in the DSM manual, and “psychoanalysis” for the treatment of mental illness. *See, e.g.*, STUART A. KIRK & HERB KUTCHINS, *THE SELLING OF DSM 7*, 11 (1992) (arguing that psychologists “worried that the DSM-III was an attempt by psychiatrists to medicalize more human problems, laying claim by professional territory that was being hotly contested by them and others.”).

122. *See* Abbott, *supra* note 108, at 72 (“The direct creation of subordinate groups has great advantages for the professions with full jurisdiction. It enables extension of dominant effort without division of dominant prerequisites.”).

to render services under the dominant expert community's supervision.¹²³ For example, doctors have delegated the provision of on-site emergency aid (or pre-hospital aid) to paramedics, whose conduct is regulated by, for example, the "Basic Life Support Guidelines" and "Advanced Life Support Guidelines," both codifications of doctors' abstract knowledge.¹²⁴ As a general rule, deviation from these guidelines requires direct medical oversight, thus sharply reducing paramedic discretion.¹²⁵

Codification of expert knowledge, however, also makes knowledge more accessible to non-specialists and would be expected to ultimately erode specialists' control over that knowledge domain. Indeed, some medical sociologists predict that the use of information systems to monitor medical examinations, assist with diagnoses, and direct treatment plans would lead to a considerable erosion of physicians' power and autonomy.¹²⁶ But just as complete codification of expert knowledge would erode jurisdictional control, so would absolute abstraction. Abstract knowledge that remains completely inaccessible to the lay public precludes the public from evaluating the effectiveness of the expert community's jurisdictional claims, especially if that community also controls the tests that evaluate effectiveness itself. Absolute abstraction demands absolute trust in individual members of the profession as possessing the required, inaccessible expertise to solve the relevant problems. Therefore, codification also serves to legitimate a claim to effective treatment by making accessible to the lay public an expert community's claims to expertise through a simplified, rule-based version of

123. This particular role for rules as devices to control subordinate communities aligns most closely with traditional principal-agent models of judicial decision-making, in which principals use rules to constrain their agents' discretion. *See, e.g.,* Tonja Jacobi & Emerson H. Tiller, *Legal Doctrine and Political Control*, 23 J.L. ECON. & ORG. 326, 339 (2007) (modeling a higher court's choice of clear rules or flexible standards on the level of political alignment between the two courts, with higher political alignment resulting in a choice of standards and vice versa).

124. *See, e.g.,* Jerry P. Nolan et al., ERC Guidelines Writing Group, *European Resuscitation Council Guidelines for Resuscitation 2010*, 81 RESUSCITATION 1219–1276 (2010); STAMFORD HOSP. EMS SPONSOR HOSP. PROGRAM, ADVANCED CLINICAL CARE GUIDELINES, POLICIES AND PROCEDURES, 1 (2013), available at <http://www.stamfordems.org/Library/2012%20Adult%20and%20Pedi%20Treatment%20Guidelines.pdf>.

125. Stamford Hospital, *supra* note 124, at 4 ("No provider shall deviate from these guidelines without direct medical oversight . . .").

126. *See, e.g.,* John D. Stoeckle, *Reflections on Modern Doctoring*, 66 MILLBANK Q. 76, 82–83 (1988). Similarly, several legal scholars predict that information systems, allowing for the proliferation of low-cost, do-it-yourself court filings, would erode lawyers' control over the provision of legal services. *See, e.g.,* Richard L. Marcus, *The Impact of Computers on the Legal Profession: Evolution or Revolution?*, 102 NW. U. L. REV. 1827 (2008) (listing sources).

the experts' knowledge.¹²⁷ Thus, jurisdictional control requires a balance between codification and abstraction.

For example, sociologists of medicine argue that the DSM—a manual that codifies mental health diagnostic categories—was developed to address “the [psychiatric] profession’s self-doubts and its vulnerability to public and scientific criticism.”¹²⁸ Similarly, the turn towards the standardization of medical care was made, at least in part, to address a legitimacy crisis.¹²⁹ A series of studies documented wide divergences in the medical treatments offered to similarly situated patients, thus undermining the credibility of health care practitioners.¹³⁰ The solution adopted by the medical profession was to convene expert medical panels to draft rule-like clinical practice guidelines for a range of medical procedures, based on the best available scientific evidence.¹³¹ In both of these cases, however, codification reduced the discretion of members *within* the expert group, not just the discretion of members of subordinate professions. Because the adoption of guidelines limited doctors’ ability to rely on their intuition based on practice experience, they have proved controversial.¹³² Indeed, several studies have found that only a minority of doctors complies with guidelines in their field, despite being familiar with them.¹³³

Because codification serves not only to control subordinate communities, but also to restore the public’s trust in the objectivity and reliability of expert judgment, expert communities that enjoy a low level of public trust are expected to rely on inflexible rules most often.¹³⁴

127. See, e.g., THEODORE M. PORTER, TRUST IN NUMBERS: THE PURSUIT OF OBJECTIVITY IN SCIENCE AND PUBLIC LIFE 4 (1995) (“Mechanical objectivity . . . has a powerful appeal to the wider public. It implies personal restraint. It means following the rules.”); Abbott *supra* note 108, at 60 (“By revealing to the public some of its professional terminology and insights, a profession attracts public sympathy to its own definition of tasks and its own approach to solving them.”).

128. Kirk & Kutchins, *supra* note 121, at 13.

129. See, e.g., Stefan Timmermans & Emily S. Kolker, *Evidence-Based Medicine and the Reconfiguration of Medical Knowledge*, 45 J. HEALTH & SOC. BEHAV., 177, 177 (2004).

130. See, e.g., THE CTR. FOR THE EVALUATIVE CLINICAL SCI., DARTMOUTH MEDICAL SCHOOL, THE DARTMOUTH ATLAS OF HEALTH CARE 1998 2 (John E. Wennberg et al. eds., 1998) (mapping the frequency and variety of surgical interventions by geographical area to similarly situated patients, and showing that “in health care, geography is destiny”).

131. See Timmermans & Kolker, *supra* note 129, at 181.

132. *Id.* at 186.

133. *Id.* (noting that other studies have “overestimated the role of guidelines in daily clinical practice”).

134. Porter, *supra* note 127, at 229–30 (noting that recourse to inflexible rules is most salient in scientific communities that are poorly insulated from public criticism).

B. THE EMBODIED APPROACH TO EXPERTISE: HOW NOVICES BECOME EXPERTS

A second strand of sociological thought has focused on the interplay between explicit rules of decision and contextual knowledge in both expert training and expert decision-making.¹³⁵ The key insight derived from this line of inquiry is that expertise is inextricably linked with tacit knowledge—“inarticulatable skills of which one cannot fully give account”¹³⁶—that make it impossible to fully codify an expert’s body of knowledge into a set of written rules.

Nevertheless, rule making plays an important role in accounts of expertise acquisition. Self-conscious following of explicit rules is what enables a novice to begin his or her path towards expertise.¹³⁷ But while the novice applies “context free” rules—being incapable of taking into account contextual factors that may require the modification of these rules—an expert not only internalizes but also transcends rules.¹³⁸ While a novice slowly and deliberately strives to follow rules, through a “painful” and “jerky” process, an expert experiences “flow” as he “unselfconsciously” recognizes complex contextual cues.¹³⁹ In fact, experts often tend not to follow the heuristics they relied upon during their training.¹⁴⁰ An expert relates to context in “a fluid way using cues that it is impossible to articulate and that if articulated would usually not correspond, or might even contradict, the rules explained to novices.”¹⁴¹ Thus, a direct consequence of expert intuition is an unavoidable conflict between the rules as explained to novices and their actual application by experts to real-world contexts.

Gaining expertise, however, requires more than following rules that are eventually transcended through repeated practice. Rather, to fully grasp expert knowledge requires “enculturation”: interactively immersing oneself

135. See, e.g., HARRY COLLINS, TACIT AND EXPLICIT KNOWLEDGE (2010); HARRY COLLINS & ROBERT EVANS, RETHINKING EXPERTISE (2007); HUBERT L. DREYFUS & STUART E. DREYFUS, MIND OVER MACHINE: THE POWER OF HUMAN INTUITION AND EXPERTISE IN THE ERA OF THE COMPUTER (1986); Robin Cowan, *Expert Systems: Aspects of and Limitations to the Codifiability of Knowledge*, 30 RES. POL. 1355, 1356 (2001) (describing the limitations of a computer expert system—i.e., computer code designed to simulate expert decision-making—as lacking the ability to make contextual decisions).

136. EVAN SELINGER, EXPERTISE: PHILOSOPHICAL REFLECTIONS 10 (2011).

137. Dreyfus & Dreyfus, *supra* note 135, at 21–36.

138. *Id.*

139. *Id.*

140. Selinger, *supra* note 136, at 19 (describing experts as having “acquired and embodied skills that provide the basis for determining whether rule following or intuitive comportment are meaningful guides for acting in the field one becomes expert in”).

141. Collins & Evans, *supra* note 135, at 25.

into expert culture.¹⁴² In other words, acquiring expertise requires learning by doing.¹⁴³ One important consequence of locating expertise within the expert community rather than with the individual is that both becoming and continuing to be an “expert” requires embeddedness in the relevant expert community: “expertise can be lost if time is spent away from the group.”¹⁴⁴

Despite this division between novices and experts—with the former relying more heavily on rules to make decisions—there is a high level of variability in the extent to which experts themselves resort to rules as decision-making tools. Theodore Porter’s research on expert communities in science shows that communities that are fractured by internal dissent and that lack widely shared background assumptions—or what he terms “weak” communities—tend to be more rule-bound than established, “strong” expert communities.¹⁴⁵ Experts in strong communities are more likely to rely on tacit, contextual knowledge to make decisions, and to relegate rules to novice training.¹⁴⁶ For example, Porter describes how the research community of high-energy physicists relies on the “long process of socialization” and a “tight network of personal contacts” to operate with a high degree of informality. Only graduate students (that is, novices) “pay much attention to published papers; mature scientists interact mainly by talking, not writing.”¹⁴⁷

Finally, from these studies emerge two additional insights. First, when beginners reach the expert stage they are transformed, not only in their ability to dispense with rules, but also in their affective relationship to their field of expertise.¹⁴⁸ The process of acquiring expertise represents a progression “from relative detachment to engaged commitment.”¹⁴⁹ Second, experts will have difficulty communicating with non-experts precisely because non-

142. Collins & Evans, *supra* note 135, at 23–24.

143. *See, e.g.*, THOMAS KUHN, *STRUCTURE OF SCIENTIFIC REVOLUTIONS* 47 (1962) (pointing out that serious science is done only by those who have been well socialized into a body of specialists).

144. Collins & Evans, *supra* note 135, at 3.

145. Porter, *supra* note 127, at 228 (“[T]he relative rigidity of rules for composing papers, analyzing data, even formulating theory, ought to be understood in part as a way of generating shared discourse, of unifying a weak research community.”).

146. *Id.* at 230 (“In . . . fields dominated by a relatively secure community, much of what we normally associate with the scientific mentality—such as an insistence on objectivity, on the written word, on rigorous quantification—is to a surprising degree missing. Scientific knowledge is most likely to display conspicuously the trappings of science in fields with insecure borders, communities with persistent boundary problems.”).

147. *Id.* at 222.

148. Dreyfus & Dreyfus, *supra* note 135, at 19 (arguing that it is this affective transformation following skill acquisition—in addition to tacit knowledge—that differentiates expert communities from computerized expert systems.).

149. *See id.* at 35.

experts can be likened to novices who only have access to the rules, but not the intuition, of the expert community.¹⁵⁰

IV. A TYPOLOGY OF FEATURES OF EXPERT COMMUNITIES

Understanding the behavior of organized groups of experts requires recognizing their unappreciated dynamic of jurisdictional competition for maximal control and autonomy in the development and application of an expert body of knowledge. It also requires an appreciation of the mechanisms through which that competition takes place, including the multiple roles of formal rules as mechanisms of controlling subordinate communities and increasing external legitimacy, as well as the importance of competing framings of tasks and problems. Finally, it requires understanding how expert decision-making differs from that of novices. Applying these insights to the Federal Circuit and potentially to specialized courts more broadly, however, involves, first, an exploration of whether the Federal Circuit can be analogized to the expert communities studied by sociologists of expertise and described in Part III. Doing so requires a fine-grained understanding of how various aspects of expert community behavior interact with each other. Second, these insights must be adapted to the hierarchical court system. Drawing from and synthesizing the work discussed in Part III, Section IV.A argues that the Federal Circuit behaves like an expert community. Section IV.B unbundles expert behavior into four inter-related characteristics: (1) epistemic control, (2) codification, (3) typecasting, and (4) inability to self-coordinate. Finally, Section IV.C discusses alternative explanations of Federal Circuit behavior.

A. THE FEDERAL CIRCUIT AS AN EXPERT COMMUNITY

Much of the research presented in the previous section studied traditional professional groups—institutions that are largely autonomous from the state, with independent entrance exams, licensing procedures, and ethical guidelines. Does this research apply to an institution like the Federal Circuit? Answering this question requires understanding both similarities and differences between the Federal Circuit and those expert groups studied by sociologists. Where the Federal Circuit is different, how might it be different, and how might these differences impact behavior predictions that flow from these sociological studies?

150. See Selinger, *supra* note 136, at 22.

Sociologists who study traditional professions already have in mind a broader definition of the term “profession” than how the term is colloquially understood. For example, sociologist Andrew Abbott adopted what he termed a “very loose” definition of “profession” in his work as “somewhat exclusive groups of individuals applying somewhat abstract knowledge to particular cases.”¹⁵¹ And sociologist Gil Eyal argues that jurisdictional competition can take place between “any groups that can lay a claim of expertise.”¹⁵² In turn, this suggests that insights derived from the sociology of expertise are applicable to the Federal Circuit, so long as it can be conceptualized as a relatively exclusive group with a claim to expertise in patent law.

Expertise in patent law can be disaggregated at least into three levels. First, the Federal Circuit has particular expertise in formulating patent doctrine to fulfill the dual congressional mandate of uniformity and efficiency.¹⁵³ Second, the Federal Circuit has special knowledge on how to apply abstract patent doctrine to technical fact patterns.¹⁵⁴ Third, the Federal Circuit has technical expertise, which involves an understanding (or at least a comparative advantage vis-à-vis other courts) of the complex and evolving technology often involved in patent litigation.¹⁵⁵ Although it is debatable whether the Federal Circuit judges *in fact* possess the required legal and technical expertise, Congress, other courts (including the Supreme Court), and academic commentators have *attributed*¹⁵⁶ these types of expertise to the Federal Circuit.¹⁵⁷ Importantly, recent empirical studies show that district

151. Abbott, *supra* note 108, at 318.

152. Gil Eyal, *For a Sociology of Expertise: The Social Origins of the Autism Epidemic*, 118 AM. J. SOC. 863, 869 (2013). See Gorman & Sandefur, *supra* note 108, at 277 (arguing that “the commonalities between traditional professions and new forms of knowledge-based work [were] more important than their differences.”).

153. See H.R. Rep. No. 97-312, at 20 (1981).

154. See *id.* at 22–23; S. Rep. No. 97-275, at 23 (1981).

155. For example, Daniel Meador, a key figure in the creation of the Federal Circuit, praised its limited subject matter jurisdiction as a means to increase the judges’ expertise in that subject area. See S. Rep. No. 97-275, at 6 (1981).

156. Attributed, as opposed to substantive, expertise does not require an objective, external measure of competence. Rather, under an attributional view of expertise, being an expert means that other relevant audiences attribute the quality of expertise to a particular community that also views itself as expert.

157. See, e.g., *Warner-Jenkinson Co. v. Hilton Davis Chemical Co.*, 520 U.S. 17, 40 (1997) (noting the Federal Circuit’s “special expertise” in patent law); Transcript of Oral Argument at 28, *Alice Corp. v. CLS Bank Int’l*, 134 S. Ct. 2347 (2014) (No. 13-298) (Breyer, J.) (describing the Supreme Court decisions in *Mayo*, 132 S. Ct. 1289 and *Bilski v. Kappos*, 561 U.S. 593 (2010) as “sketch[ing] an outer shell of the content, hoping that the experts, you and the other lawyers and the . . . circuit court, could fill in a little better than we had done the content of that shell”); John M. Golden, *supra* note 157, at 660 (arguing that the

courts accord the Federal Circuit greater institutional authority in patent law (compared to the Supreme Court) than they accord other circuit courts in copyright law (compared to the Supreme Court).¹⁵⁸ This suggests that district courts also view the Federal Circuit as deserving of increased deference in issues of patent law, likely by virtue of their relevant expertise. In addition, the Federal Circuit has self-identified as an expert community. For example, the Federal Circuit has often noted that it possesses “special” and “useful” expertise on matters of patent law based in part on the large volume of patent cases it decides.¹⁵⁹

There are, however, two key differences between the Federal Circuit and the expert groups described in the previous sections: (1) embeddedness in a hierarchical court structure and (2) high epistemic diversity among members of the court. Because the Federal Circuit is embedded in a hierarchical court structure, it is subject to rules of deference (such as Rule 52(a) requiring deference to the district court’s findings of fact, or the required deference to Supreme Court holdings) that place limits on its autonomy and thus on its ability to compete freely with other expert communities. In addition, current Federal Circuit judges have no say over new judicial appointments¹⁶⁰—in contrast to most expert groups that control admittance into their community. Of course, other expert communities are subject to external controls as well. For example, medical malpractice law regulates doctors’ behavior.¹⁶¹ And several states have passed laws banning some types of scientific research, such as human reproductive and therapeutic cloning.¹⁶² Nevertheless, the hierarchical structure of the judicial system does not map neatly onto other types of regulation imposed on the expert communities traditionally studied by sociologists.

Federal Circuit “hears enough patent cases to acquire unquestionable expertise on questions of substantive patent law”).

158. David R. Pekarek Krohn & Emerson H. Tiller, *Federal Circuit Patent Precedent: An Empirical Study of Institutional Authority and Intellectual Property Ideology*, 2012 WIS. L. REV. 1177, 1194 (2012).

159. *See Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 234 F.3d 558, 571–72 (Fed. Cir. 2000) (en banc). *See also Highmark, Inc. v. Allcare Health Mgmt. Sys., Inc.*, 687 F.3d 1300 (Fed. Cir. 2012), *reh’g denied*, 701 F.3d 1351, 1356 (Fed. Cir. 2012) (“The Federal Circuit brings to the table useful expertise. Our court sees far more patent cases than any district court, and is well positioned to recognize those ‘exceptional’ cases in which a litigant could not, under the law, have had a reasonable expectation of success.”).

160. *See* U.S. CONST. art. II, § 2, cl. 2; 28 U.S.C. § 44(a) (2006).

161. 42 U.S.C. § 11131 (2006).

162. *See, e.g.*, ARK. CODE ANN. §§ 20-16-1001-1004 (2003); 2005 Ind. Acts No. 268; MICH. COMP. LAWS §§ 333.26401-26403 (1999); N.D. CENT. CODE §§ 12.1-39 (2003); S.D. CODIFIED LAWS § 34-14-27 (2004); VA. CODE ANN. §§ 32.1-162.22-2 (2001).

Second, there is greater diversity in the background and training of Federal Circuit judges (what I term “epistemic diversity”) than would be expected of members of traditional professions. There is no standard curriculum that makes a Federal Circuit judge an expert in patent law.

This level of epistemic diversity among Federal Circuit judges is not completely unheard of in other expert communities. In fact, many expert communities could be subdivided into sub-communities with closer epistemic connections. For example, expert communities of geneticists contain within them communities of human geneticists, mouse geneticists, fruit fly geneticists, and so on. Nevertheless, the epistemic diversity of the Federal Circuit raises the question of whether it can still be considered a coherent, single community. Several lines of argument indicate that the Federal Circuit does indeed behave like a single expert community—albeit one with a high potential for internal dissent and fracture. First, the Federal Circuit sees itself as an institution with a collective “special expertise” in patent law—a self-perception that is shared across government actors, including Congress, lower courts, and the Supreme Court.¹⁶³ Second, epistemic diversity upon entering the court does not preclude the development of shared norms in the course of making patent decisions. Indeed, an important finding of sociologists of expertise is that communities that work towards a shared goal (in this case, to develop a coherent body of patent law) will tend to develop shared understandings and norms.¹⁶⁴ In this framework, new members of the Federal Circuit are expected to be enculturated into existing Federal Circuit norms. Still, high levels of epistemic diversity are likely to make the Federal Circuit more akin to “weak” expert communities, with high levels of internal dissent.

Finally, the Federal Circuit has recently experienced a high turnover rate in its judicial membership. Six of the eleven active Federal Circuit judges have been appointed by the Obama administration between 2010 and 2013, imposing a potentially important disruption of the CAFC’s composition.¹⁶⁵

163. See *supra* notes 164–166.

164. See, e.g., Thomas Gieryn, *Boundaries of Science*, in *THE HANDBOOK OF SCIENCE AND TECHNOLOGY STUDIES* 393, 412 (Sheila Jasanoff, Gerald E. Markle, James C. Petersen & Trevor Pinch, eds., 1995).

165. Kathleen M. O’Malley, *Circuit Judge*, UNITED STATES COURT OF APPEALS FOR THE FEDERAL CIRCUIT, <http://www.cafc.uscourts.gov/judges/kathleen-m-omalley-circuit-judge.html>; Jimmie V. Reyna, *Circuit Judge*, UNITED STATES COURT OF APPEALS FOR THE FEDERAL CIRCUIT, <http://www.cafc.uscourts.gov/judges/jimmie-v-reyna-circuit-judge.html>; Evan J. Wallach, *Circuit Judge*, UNITED STATES COURT OF APPEALS FOR THE FEDERAL CIRCUIT, <http://www.cafc.uscourts.gov/judges/evan-j-wallach-circuit-judge.html>; Richard G. Taranto, *Circuit Judge*, UNITED STATES COURT OF APPEALS FOR THE FEDERAL

Can a high-turnover, small group of judges be analogized to an “expert community” of the type studied by sociologists of expertise? It is certainly the case that one potential hurdle to conceptualizing the Federal Circuit as an expert community concerns the stability of its expertise given judge turnover. The model of expert communities introduced here assumes the development of a set of practices and social norms associated with the Federal Circuit’s role as an expert in patent law.¹⁶⁶ It further assumes that those norms are likely to persist in spite of personnel changes, as new judges are enculturated into the practices of the expert community.¹⁶⁷ Of course, this logic has a limit: if all Federal Circuit judges were to be replaced overnight, there would remain no or little institutional repository for existing norms. But given a stable core of Federal Circuit judges over time the assumption of a stable set of practices is likely justified. There may be particular times in the Federal Circuit’s history, however, when its institutional identity as an expert court is particularly fragile. During these times, we would expect the Federal Circuit to mirror the behavior of “weak” expert communities described in Part III.

The typology of Federal Circuit decision-making developed below adapts insights from the sociology of expertise to the context of specialized courts in light of the hierarchical structure in which the court is embedded and the epistemic diversity of the Federal Circuit.

B. FOUR FEATURES OF EXPERT COMMUNITIES AND THEIR APPLICATION TO FEDERAL CIRCUIT BEHAVIOR

This Section introduces the four features of the typology and their interrelationship. It then applies each feature to the Federal Circuit, showing how conceptualizing the Federal Circuit as an expert community has both explanatory and predictive power for Federal Circuit behavior.

1. *Epistemic Control*

Jurisdictional control over both abstract analyses and solutions for problems within an expert community’s field of expertise requires the twin forces of monopoly and autonomy—which this Article bundles under the term “epistemic control.” *Epistemic monopoly* allows control over the *supply* of expertise by placing the expert community as the only source of valid

CIRCUIT, <http://www.cafc.uscourts.gov/judges/richard-g-taranto-circuit-judge.html>; Raymond T. Chen, UNITED STATES COURT OF APPEALS FOR THE FEDERAL CIRCUIT, <http://www.cafc.uscourts.gov/judges/raymond-t-chen-circuit-judge.html>; Todd M. Hughes, Circuit Judge, UNITED STATES COURT OF APPEALS FOR THE FEDERAL CIRCUIT, <http://www.cafc.uscourts.gov/judges/todd-m-hughes-circuit-judge.html>.

166. See, e.g., Pedraza-Fariña, *supra* note 91, at 844–45.

167. *Id.*

solutions for a particular problem. *Epistemic autonomy* refers to an expert community's independence in defining the significance and relevance of its knowledge base. Autonomy leads to jurisdictional control over the classification and definition of a problem as pertaining to an expert community's sphere of expertise. Epistemological autonomy allows control over the *demand* for expertise by granting an expert community's independence in framing its knowledge base.

An example of how these two forces may be unlinked is illustrative: a government agency could grant *epistemological monopoly* to an expert community to solve problem X but retain *epistemological autonomy* to define precisely what X is and whether X requires the application of the knowledge base of a particular community. Complete epistemic control implies control both over the system of knowledge (abstractions) used to solve a particular problem, and the framing of the problem itself as amenable to solution by that particular set of abstractions.

Several consequences follow from an expert community's drive for epistemic control. First, epistemic monopoly will reduce deference to findings by subordinate non-experts—thus helping *explain* the CAFC's non-deferential relationship to lower courts (which act as subordinate generalist communities). Second, the drive for epistemic monopoly will lead to jurisdictional competition with other expert communities—thus helping *explain* the CAFC's fraught relationship with the PTO, characterized by both lack of deference on the CAFC's part and by the PTO's resistance and defiance to the CAFC. Third, epistemic monopoly *predicts* resistance to solutions for a particular problem proposed by non-experts in positions of authority—the problem of *defiance*. Finally, epistemic autonomy is likely to lead to resistance to alternative framings of or solutions for the problems under study that usurp an expert community's ability to address that problem—thus predicting *jurisdictional expansion*. These four implications of epistemic control are discussed more fully below.

a) Explaining the Federal Circuit's Relationship with Lower Courts:
Non-Deference

As discussed in Section II.A, the Federal Circuit has arrogated power over facts, or construed mixed questions of law and fact as questions of law—issues on which appellate courts traditionally grant deference to district courts. This model complements existing explanations of Federal Circuit behavior that attribute its lack of deference to lower courts simply to its

better judgment on or knowledge of patent issues.¹⁶⁸ If superior understanding of how to apply patent law to a particular technological area were the only factor driving this lack of deference, one may expect the Federal Circuit to follow the general rule of deference to trial courts on factual matters but make reasoned, case-by-case corrections when its understanding of the technology or reliability of expert testimony differed from that of the district court. Instead, the Federal Circuit has resorted to blanket rules of non- or reduced deference that increase its monopoly on decisional authority, and allow it to avoid having to give explanations for deviating from a trial court's interpretation of expert testimony. Indeed, it is telling that the Federal Circuit has been most resistant to show deference to district courts on claim construction—an issue that is often outcome-determinative of all other questions in a patent case.¹⁶⁹ This behavior is more consistent with an expert community's drive for maximal control over the supply of expertise in patent law.

The Federal Circuit's lack of deference to trial courts on issues of patent law stands in sharp contrast to the court's non-patent decisions, which are characterized by high affirmance rates and deferential standards of review.¹⁷⁰ Patent law scholars have relied on theories of institutional design of administrative agencies to explain the Federal Circuit's disparate treatment of patent and non-patent cases.¹⁷¹ Specifically, these theories predict that agencies with multiple tasks will tend to give prominence to one of those tasks that implicate “agency culture, history, monitoring difficulties, and political concerns.”¹⁷² Under this view, the Federal Circuit expands its jurisdiction on issues central to its core mission, but surrenders it on peripheral issues.¹⁷³ This model offers an alternative, yet complementary,

168. See, e.g., Dreyfuss, *supra* note 40, at 48 (“A trial judge who has never read a technical document before is less likely to interpret it correctly, no matter how many expert witnesses are called to testify, than an appellate judge who has extensive experience in dealing with such matters.”).

169. See, e.g., Oskar Liivak, *Rescuing the Invention from the Cult of the Claim*, 42 SETON HALL L. REV. 1, 7 (2012) (“[E]very substantive aspect of patent law is controlled by the claims.”); Giles S. Rich, *Extent of Protection and Interpretation of Claims—American Perspectives*, 21 INT'L REV. INDUS. PROP. & COPYRIGHT L. 497, 499 (1990) (“The claim is the name of the game.”).

170. See Ted L. Field, “Judicial Hyperactivity” in the Federal Circuit: *An Empirical Study*, 46 U.S.F. L. REV. 721, 759–63 (2012) (reporting a 28.8% reversal rate in patent cases and a 14.3% reversal rate in nonpatent cases).

171. See, e.g., Paul R. Gugliuzza, *The Federal Circuit As A Federal Court*, 54 WM. & MARY L. REV. 1791 (2013).

172. *Id.* at 1799.

173. *Id.*

explanation: the Federal Circuit's unique behavior in the area of patent law flows from its attributed expertise in the subject.

b) Explaining the Federal Circuit's Relationship with Other Expert Bodies: Jurisdictional Competition

The drive to maintain maximal control and to displace alternative framings of a problem predicts competition with other expert communities. In the area of patent law, the Federal Circuit, the PTO, and the ITC would be expected to compete with each other for maximal monopoly and autonomy in the design and application of patent law. Competition between expert communities can take different forms, depending on the tools available to those communities to maintain and expand their epistemic monopoly and autonomy. As elaborated in the next section, codification (or rulemaking) remains one of the most powerful mechanisms where a superior expert community can both delegate authority to a subordinate expert community and control how that authority is exercised. The role of codification, and its impact on the relationship between the Federal Circuit and the PTO, is explored in depth in the next section.

But codification is not the only means of competition. The Federal Circuit has used additional strategies to avoid according deference to the PTO—from refusing to recognize the existence of factual disputes to applying a more stringent standard of review than that mandated by the APA.¹⁷⁴ On its part, the PTO has been keen on expanding its influence over patent law and policy, challenging the Federal Circuit's power at the Supreme Court and, more quietly, simply refusing to apply Federal Circuit guidelines—providing further evidence of the competitive relationship between these two communities.¹⁷⁵

c) Predicting Federal Circuit Behavior: Defying the Supreme Court

An expert community's drive to control both the system of knowledge (abstractions) used to solve a particular problem and how that problem is framed leads to an inevitable confrontation with other communities that are granted authority over the expert community. This prediction is consistent

174. See *Brand v. Miller*, 487 F.3d 862, 869 n.3 (Fed. Cir. 2007).

175. See, e.g., Clarisa Long, *The PTO and the Market for Influence in Patent Law*, 157 U. PA. L. REV. 1965, 1966 (2009) (“The PTO has been vying to gain more influence in the market for supplying legal rules and norms.”); Craig Allen Nard, *Deference, Defiance and the Useful Arts*, 56 OHIO ST. L.J. 1415 (1995) (“[T]he PTO, as of late, has displayed an independent temperament, at times to the point of defiance, and has argued for greater deference with respect to its patentability decisions and interpretations of various provisions of the patent code.”).

with findings that in hierarchical organizations (be it private or public) specialized subordinates are likely to grant less authority to their generalist superiors than are generalist subordinates.¹⁷⁶

There have been no quantitative empirical studies comparing Federal Circuit disobedience of Supreme Court decisions to disobedience by other circuits,¹⁷⁷ or assessing whether the Federal Circuit is more likely to defy the Supreme Court in its attributed area of expertise (patent law) than in any of the other cases that make up its docket—both of which are predicted by this model. Nevertheless, qualitative evidence suggests that this is the case. Indeed, Chief Justice Roberts has remarked on the Federal Circuit's unusual behavior, noting that it seemed an exception to the rule that lower courts generally follow Supreme Court precedent.¹⁷⁸ An analysis of Federal Circuit case law reveals a pattern of resistance to implementing Supreme Court decisions overruling Federal Circuit precedent—a pattern consistent with the model's prediction of defiance to decisions by generalist superiors.¹⁷⁹

For example, in one of the two Supreme Court cases reviewing the Federal Circuit in the 1980s, *Dennison Manufacturing Co. v. Panduit Corp.*, the Court remanded the case to the Federal Circuit with explicit instructions to provide an opinion “clearly setting forth” its reasoning on why Rule 52(a) did not mandate deference to the district court's factual determinations on nonobviousness.¹⁸⁰ Following the Court's decision, however, several Federal Circuit cases simply continued applying a de novo standard of review to the entire non-obviousness determination.¹⁸¹

176. JAMES Q. WILSON, BUREAUCRACY: WHAT GOVERNMENT AGENCIES DO AND WHY THEY DO IT 91–101 (1989).

177. A citation study of the Court of Customs and Patent Appeals, which existed alongside appellate courts prior to the creation of the Federal Circuit and heard appeals from PTO denials found that it “consistently cited the Supreme Court at lower rates than did the courts of appeals.” Lawrence Baum, *Specialization and Authority Acceptance: The Supreme Court and Lower Federal Courts*, 47 POL. RES. Q. 693, 700 (1994).

178. *See* *Carlsbad Tech., Inc. v. HIF Bio, Inc.*, 129 S. Ct. 1862, 1868 (2009).

179. *See supra* Part II.

180. 475 U.S. 809 (1986), *see also* Rai, *Engaging Facts and Policy*, *supra* note 8, at 1061 (analyzing *Dennison* and its aftermath).

181. *See, e.g.*, *Newell Cos. v. Kenney Mfg. Co.*, 864 F.2d 757, 765 (Fed. Cir. 1988) (arguing that the case presented no issues of fact and reviewing the trial court's non-obviousness determination de novo, even though deciding the case required solving factual dispute regarding the differences between the prior art and the patent at issue); *Uniroyal, Inc. v. Rudkin-Wiley Corp.*, 837 F.2d 1044, 1051 (Fed. Cir. 1988) (granting no deference to the trial court's findings, which were based on its evaluation of expert testimony, that a PHOSITA would have been motivated to combine two prior art references to make the invention at issue).

As discussed in Section II.A.2, in *Dickinson v. Zurko*,¹⁸² a case involving deference to the PTO's findings of fact, the Supreme Court reversed a long line of Federal Circuit precedent holding that the deferential standard of review to agency fact-finding set forth in the APA did not apply to the PTO.¹⁸³ Subsequent cases, however, continued to review PTO fact-finding more stringently than required by the APA. They did so by interpreting the APA's "substantial evidence" standard as being more stringent than the "arbitrary and capricious" standard when applied to judicial review of agency fact-finding—in contravention of Supreme Court precedent.¹⁸⁴

The Federal Circuit's tendency to stray from Supreme Court opinions extends further than cases concerning the proper standard of review of district court and agency action. As mentioned in Section II.B, in *KSR v. Teleflex*, the Supreme Court rejected the Federal Circuit's "teaching, suggestion, and motivation test" as the sole rule to determine whether an invention is "obvious" under section 103 of the Patent Act.¹⁸⁵ The Court deemed the Federal Circuit's "rigid approach" at odds with Supreme Court precedent in *Graham v. John Deere Co. of Kansas City*,¹⁸⁶ which called for a flexible, functionalist inquiry.¹⁸⁷ The Court also made clear that a real-life PHOSITA's research would not be limited by explicit teachings or suggestions to combine elements from her own field of discovery.¹⁸⁸ Rather, a PHOSITA would be driven by "design incentives and other market forces" to find solutions to existing problems worked out within the PHOSITA's own field or a different one.¹⁸⁹

KSR had clear implications for the doctrine of analogous arts, which seeks to identify the content of all relevant prior art that would be available to a PHOSITA at the time of invention.¹⁹⁰ At a minimum, it suggested that determining the contours of analogous art requires a case-by-case determination of which sources a PHOSITA would be driven to consult, given existing market forces and design incentives. Nevertheless, and despite dicta in Federal Circuit opinions recognizing that *KSR* modified the

182. 527 U.S. at 154 (1999).

183. *Id.*

184. *See, e.g.*, *Flo Healthcare Solutions, LLC v. Kappos*, 697 F.3d 1367, 1378 (Fed. Cir. 2012); *Taylor v. U.S. Patent & Trademark Office*, 339 Fed. App'x 995, 998 (Fed. Cir. 2009); *Millman v. U.S. Patent & Trademark Office*, 257 Fed. App'x 307, 311 (Fed. Cir. 2007).

185. 550 U.S. at 400.

186. 383 U.S. 1, 21 (1966).

187. *KSR*, 550 U.S. at 401.

188. *Id.* at 401–02.

189. *Id.* at 401.

190. *See* 35 U.S.C. § 112(f) (2006); *Festo*, 234 F.3d at 564.

analogous art inquiry,¹⁹¹ the Federal Circuit has adopted a formalistic approach. In fact, in an important analogous art decision announcing a new rule for determining the contours of analogous art, *In re Klein*,¹⁹² the Federal Circuit did not even cite *KSR* as relevant authority.¹⁹³

Other recent cases reflect a similar tendency to disregard Supreme Court decisions that strike down long-standing Federal Circuit doctrine. In patentable subject matter, the Federal Circuit all but ignored the Court's instructions on remand in *Association for Molecular Pathology v. Myriad Genetics*¹⁹⁴ to decide the case in light of the Court's decision in *Mayo v. Prometheus*.¹⁹⁵ *Myriad* concerned the patentability of isolated genomic DNA (i.e., DNA extracted from a cell) and cDNA (i.e., the portion of DNA that codes for a protein, which is manufactured in the laboratory). Relying on its "product of nature" doctrine, the Federal Circuit had reasoned in *Myriad* that both genomic and cDNA were patent eligible because the genomic DNA and cDNA molecules obtained by laboratory manipulation were different from those existing in their natural state inside a cell.¹⁹⁶ *Mayo* concerned the patentability of a diagnostic method for adjusting the dosage of a drug to avoid toxicity while preserving therapeutic effectiveness.¹⁹⁷ The method relied upon a finding that concentrations in the blood above a threshold level of certain drug metabolites led to toxicity. The Court in *Mayo* reasoned that the "relationships between the concentration in the blood of certain thiopurine metabolites and the likelihood that the drug dosage will be ineffective or induce harmful side-effects"¹⁹⁸ were patent-ineligible laws of nature.¹⁹⁹

An application of the reasoning in *Mayo* to *Myriad* could have led the Federal Circuit to focus on the informational content of DNA—a code that

191. *See, e.g.*, *Wyers v. Master Lock Co.*, 616 F.3d 1231, 1238 (Fed. Cir. 2010).

192. 647 F.3d 1343, 1352 (Fed. Cir. 2011).

193. *See id.*

194. *Ass'n for Molecular Pathology v. U.S. Patent and Trademark Office*, 689 F.3d 1303, 1308–09 (Fed. Cir. 2012), *aff'd in part, rev'd in part sub nom. Ass'n for Molecular Pathology v. Myriad Genetics, Inc.*, 133 S. Ct. 2107 (2013).

195. *Mayo*, 132 S. Ct. 1289.

196. More specifically, the opinion authored by Judge Lourie focused on how isolating genomic DNA required breaking chemical bonds, and how the cDNA molecule did not exist in nature, but had to be synthesized in the laboratory. *Ass'n for Molecular Pathology*, 689 F.3d at 1329.

197. *Mayo*, 132 S. Ct. at 1290–91.

198. *Id.* at 1294.

199. *Id.*

gives instructions for translating DNA into a specific protein sequence.²⁰⁰ If the DNA code is a patent-ineligible law of nature, simply separating the DNA from the genome, using what the Federal Circuit itself characterized as “routine methods,” may not have been sufficient under *Mayo* to render genomic DNA patent-eligible.²⁰¹ Nevertheless, the Federal Circuit simply declared that *Mayo* was not applicable to the issue of patentability of genomic and cDNA.²⁰²

*Gunn v. Minton*²⁰³ is another example of Federal Circuit defiance. In *Gunn*, the Supreme Court reversed the Federal Circuit’s interpretation of its own jurisdiction as encompassing virtually any state law claim that raised issues of patent validity, enforceability, or infringement.²⁰⁴ *Gunn* concerned a suit for attorney malpractice in patent litigation, filed in a Texas state court.²⁰⁵ In *Gunn*, the Supreme Court rejected the Federal Circuit’s bright-line rule that all patent malpractice cases must be brought into Federal Court.²⁰⁶ Nevertheless, even though *Gunn* rejected the Federal Circuit’s interpretation of its own jurisdictional reach,²⁰⁷ the Federal Circuit has already suggested that the decision should be interpreted narrowly, thus preserving much of its previous jurisdictional caselaw.²⁰⁸

d) Predicting Federal Circuit Behavior: Jurisdictional Expansion

The feature of epistemic autonomy predicts that the Federal Circuit will tend to frame cases that involve other bodies of law (such as state and antitrust law) but that have a patent law component, as primarily about patent law—ultimately resulting in jurisdictional expansion. An extensive scholarly literature on patent law federalism supports this prediction.²⁰⁹ For

200. The genetic code can be understood as specifying a relationship between triplets of DNA base pair molecules and single proteins. This relationship is not determined by man, but rather represents the (natural) logic that allows the reproduction of all living organisms.

201. The case involved both cDNA and genomic DNA claims. *Mayo*, 132 S. Ct. at 1294 (“We find that the process claims at issue here do not satisfy these conditions. In particular, the steps in the claimed processes (apart from the natural laws themselves) involve well-understood, routine, conventional activity previously engaged in by researchers in the field.”)

202. *Ass’n for Molecular Pathology*, 689 F.3d at 1324–25.

203. *Gunn v. Minton*, 133 S. Ct. 1059 (2013).

204. *Id.* at 1061.

205. *Id.*

206. *Id.*

207. *Id.*

208. *Forrester Env’tl Servs., Inc. v. Wheelabrator Techs., Inc.*, 715 F.3d 1329, 1334 (Fed. Cir. 2013) (noting that much of the Federal Circuit’s jurisdictional case law “may well have survived the Supreme Court’s decision in *Gunn*”).

209. See, e.g., Shubha Ghosh, *Short-Circuiting Contract Law: The Federal Circuit’s Contract Law Jurisprudence and IP Federalism*, U. Wis. Legal Studies Research, Paper No. 1247, at 2014,

example, Shubha Ghosh argues that the Federal Circuit has appropriated jurisdiction over state contract law by creating its own federal common law of contracts.²¹⁰ Similarly, the Federal Circuit has interpreted its Congressional grant of jurisdiction as encompassing any state law claim that simply requires the application of patent law.²¹¹ Paul Gugliuzza argues that this expansive interpretation is contrary to Supreme Court precedent, which granted the Federal Circuit a narrower jurisdiction over state law claims, arising only when those claims raised pure issues of patent law.²¹² And antitrust scholars have repeatedly criticized the Federal Circuit for “increasing the scope of its exclusive jurisdiction to decide appeals of antitrust and other non-patent claims that implicate issues of patent law.”²¹³

2. Codification

All communities of experts strive to codify into written rules at least part of their abstract knowledge base. Institutional approaches to expertise described in Part III focus on the *external* functions of codification—the use of codification to *teach, delegate, and control* other (often subordinate) communities, or to *gain legitimacy* in the eyes of external audiences. The institutional approach to expertise focuses on the use of rules as a tool of jurisdictional competition between expert communities. Codification enables a superior expert community to both delegate authority to a subordinate expert community and control how that authority is exercised. And by legitimizing its approach in the eyes of relevant actors, expert communities compete with each other for the demand of their services. In contrast, embodied approaches to expertise focus on the *internal* functions of codification—the use of rules for purposes internal to the expert community (to *train novices* and to *manage internal dissent*). The interplay between the internal and external functions of codification is depicted in Figure 1 below.

The codification feature provides a richer account of the role of rule-making in expert communities than traditional explanations for the choice of rules over standards that tend to focus on rules solely as a form of control.

available at <http://ssrn.com/abstract=2390214>; Gugliuzza, *supra* [Fed. Cir. as Fed. Court] note 171; Camilla A. Hrdy, *State Patent Laws in the Age of Laissez Faire*, 28 BERKELEY TECH. L.J. 45, 47 (2013); Xuan-Thao Nguyen, *Dynamic Federalism and Patent Law Reform*, 85 IND. L.J. 449 (2010).

210. Ghosh, *supra* note 209, at 3.

211. Gugliuzza, *supra* note 11, at 30.

212. *Id.* at 31–32.

213. Scott A. Stempel & John F. Terzaken III, *Casting a Long IP Shadow over Antitrust Jurisprudence: The Federal Circuit's Expanding Jurisdictional Reach*, 69 ANTITRUST L.J. 711, 711 (2002); see also Ronald S. Katz & Adam J. Safer, *In Ruling on Antitrust, Does Fed. Circuit Overstep?*, NAT'L L.J., at C20, (Oct. 16, 2000).

Specifically, it predicts that the Federal Circuit will resort to rules not only to control and teach subordinate communities, but also as a key mechanism to manage its relationships with other relevant audiences, such as the patent bar and the public at large, and to diminish discord among its own members.

These features of codification, however, need to be modified to take into account both the hierarchical structure in which the Federal Circuit is embedded and the epistemic diversity of the Federal Circuit. First, court hierarchy may introduce a paradox that is not traditionally present in other expert communities, in which members of the expert group are free to announce rules to guide subordinate—but not expert—behavior.²¹⁴ Specifically, rules designed to control or teach subordinates also bind the Federal Circuit, thus preventing it from deploying its own expertise. The Federal Circuit appears to have solved this paradox by often ignoring its own rules, a phenomenon that will be explored in the next section. Second, the high levels of epistemological diversity characteristic of the Federal Circuit, and its recent high turnover rate, suggest that it will behave like a weak expert community and attempt to gain legitimacy through codification. Finally, as a weak community, this model predicts that the Federal Circuit would resort to rules on issues characterized by high levels of internal dissent.

The next sections explore in more depth how the codification feature of expert communities can both explain and predict Federal Circuit behavior.

a) Explaining the Federal Circuit's Relationship with Lower Courts and Agencies: Teaching, Delegation, and Control

The Federal Circuit is in a dual relationship with district courts. On the one hand, the expert Federal Circuit has a *teaching relationship* with generalist district courts, which can be conceptualized as non-experts in the patent law field. As novices, district courts need rules to begin to learn the art of making patent law decisions and cannot be trusted to correctly implement standards, or open-ended, flexible inquiries.²¹⁵ On the other hand, the district court is a *subordinate community* vis-à-vis the Federal Circuit. In this context, the Federal Circuit can be expected to use codification as a means to both delegate a subset of tasks to district courts, and to tightly control the exercise of that delegation.

214. See *supra* text accompanying note 124 (explaining the relationship between paramedics and medical doctors).

215. This view is consistent with recent comments by Justice Alito during oral argument in the case *Octane Fitness*. Justice Alito suggested that the district court does not have sufficient experience in patent law to be able to judge, without more explicit guidance, whether a patent law case is “exceptional.” Transcript of Oral Argument at 13, *Octane Fitness, LLC v. ICON Health & Fitness*, 134 S. Ct. 1749 (No. 12-1184) (Alito, J.).

Examples abound of rule development by the Federal Circuit and its admonition to district courts that rules need to be rigidly implemented. For example, in *KSR* itself, the Federal Circuit chastised the district court for failing to make specific findings on what “understanding or principle within the knowledge of a skill artisan . . . would have motivated one with no knowledge of [the] invention to make the combination in the manner claimed.”²¹⁶ In other words, the Federal Circuit demanded that the district court make its reasoning explicit, which in practice meant finding prior art of record demonstrating a reason to combine references.²¹⁷ The Federal Circuit thus denied the lower court recourse to its own judgment in determining both the skill in the art in the relevant technology and whether an artisan of that skill would have combined the references under consideration. Importantly, finding the level of skill in the art, and elucidating in light of all the factual evidence whether a PHOSITA would have combined the references at issue, is a fact-laden inquiry of the type that district courts are traditionally in the best position to perform.

The boundary between these twin functions of rules—as teachers, and as instruments of delegation and control—is not sharply delineated. Using rules to teach implies controlling what tasks are delegated to novices and how those tasks are performed. The key distinction is that the *delegation* and *control* functions of codification take place in the context of a competitive, rather than a mentoring, relationship between communities. Importantly, teaching also implies relinquishing at least some control after learning has taken place.

Disentangling whether the Federal Circuit is acting as a teacher or as a delegator/controller can be quite difficult given their overlap, but one can make some testable predictions as to the likely consequences of Federal Circuit behavior in each one of these roles. First, the teaching function implies that the Federal Circuit will modify its behavior as a function of district court learning, thus relaxing control by allowing more flexibility in the application of rules. There is some evidence suggesting that the Federal Circuit has increased its deference to district court decisions,²¹⁸ and in particular to decisions by specific district court judges who have sat with the Federal Circuit,²¹⁹ which is consistent with the teaching function of

216. *Teleflex, Inc. v. KSR Int'l Co.*, 119 Fed. App'x 282, 288 (Fed. Cir. 2005).

217. *See, e.g., ACS Hosp. Sys., Inc. v. Montefiore Hosp.*, 732 F.2d 1572, 1577 (Fed. Cir. 1984); *In re Geiger*, 815 F.2d 686, 688 (Fed. Cir. 1987).

218. Anderson & Menell, *supra* note 46; Jason Rantanen & Lee Petherbridge, *Disuniformity*, Univ. Iowa Legal Studies Research, Paper No. 13–42, 2014, available at <http://ssrn.com/abstract=2351993> and <http://dx.doi.org/10.2139/ssrn.2351993>.

219. Mark A. Lemley & Shawn P. Miller, *If You Can't Beat 'Em, Join 'Em? How Sitting by Designation Affects Judicial Behavior*, Stanford Public Law Working Paper No. 2449349, available

codification.²²⁰ Second, the tension between teaching and control implies the existence of a tipping point in which greater district court expertise won't lead to greater Federal Circuit deference, because an expert district court will become in fact a competitor to the Federal Circuit.²²¹

The logic of competition, rather than mentoring, is predicted to be most prominent in the Federal Circuit's interactions with other expert communities such as the PTO and the ITC—communities that can claim to have developed their own expert abstract knowledge base in patent law and policy. Rules control the PTO or the ITC by denying them recourse to their own expertise while simultaneously cementing the Federal Circuit's epistemic control over patent policy. Indeed, the Federal Circuit has similarly relied on a rigid interpretation of the TSM test to limit the PTO's ability to use its technical expertise.²²²

The teaching function of codification is consistent with Peter Lee's information-cost theory, which argues that expert communities resort to rules to codify and simplify expert knowledge for external, lay consumption.²²³ But as shown here, an equally important function of codification is that it permits an expert community to serve as the gatekeeper of its own jurisdictional power.

Finally, the delegation/control function gives rise to an important paradox in the context of a court hierarchy. As emphasized in the previous Section, the Federal Circuit would be expected to look for ways to free itself from the very rules it created to constrain their subordinates when these rules do not accord with their own intuitions. Indeed, the Federal Circuit appears to often “break” its own rules. For example, district court judges have

at SSRN: <http://ssrn.com/abstract=2449349>. For a list of district court judges who have visited the Federal Circuit from 2006 to 2013, see <http://www.cafc.uscourts.gov/images/stories/judicial-reports/vjchartforwebsite2006-2013.pdf>.

220. Increased deference to the opinions of fellow expert community members takes place not simply because new expert members “know better” than non-members, but also because they have been socialized into the practices of the expert community and have, as a consequence, gained the trust of their peers. *See, e.g.*, Porter, *supra* note 127, at 223 (arguing that “informal judgments of character and reliability are crucial for interpreting . . . experiments”). Lemley and Miller conclude that it is this increase in trust that accounts for increased deference to district court judges who sit by designation in the Federal Circuit. *Supra* note 219, at 28.

221. *See* Arti Rai, *Specialized Trial Courts: Concentrating Expertise on Fact*, 17 BERKELEY TECH. L.J. 877 (2002).

222. *See, e.g.*, *In re Lee*, 277 F.3d 1338, 1430 (Fed. Cir. 2002) (“[C]ommon knowledge and common sense even if assumed to derive from the agency's expertise, do not substitute for authority when the law requires authority.”); *In re Rouffet*, 149 F.3d 1350, 1356–57 (Fed. Cir. 1998); *In re Fritch*, 972 F.2d 1260, 1265–66 (Fed. Cir. 1992).

223. Lee, *supra* note 7. *See infra* Part C.II.

criticized the Federal Circuit for routinely ignoring its own rules in matters of claim construction.²²⁴ Commentators have leveled a similar criticism to the Federal Circuit's choice of law jurisprudence, noting how the court "inconsistently applied its choice of law rules"²²⁵ or "simply ignore[d] the choice of law rules issue altogether."²²⁶ This paradox also makes a rule-based system much less predictable than would otherwise be anticipated. In turn, this helps explain the observation that the Federal Circuit has in fact failed to bring uniformity and predictability to its docket.²²⁷

The degree to which the Federal Circuit will in fact break its own rules when those rules do not accord with its tacit or contextual knowledge will depend, however, on the extent to which rules also serve to secure external legitimacy and manage internal dissent—a role which is more important in weak expert communities. These two additional functions require that experts themselves abide by their rules and provide clear explanations when they choose to deviate from them.

The "rules vs. standards" debate in the legal literature has not generally considered these two additional functions of rules.²²⁸ The next Section begins to fill this gap by applying these two features of codification to Federal Circuit behavior.

b) Predicting External Relationships: Seeking Legitimacy through Rules

To generate demand for their services and acceptance of their diagnoses and treatments, expert communities require sociological legitimacy from relevant audiences.²²⁹ Sociological legitimacy refers to the acceptance (by the public at large, or by specific relevant audiences) of a particular expert community's authority in its area of expertise, based on reasons other than fear of sanctions or expectations of personal gain.²³⁰ Expert communities

224. See, e.g., O'Malley, Saris & Whyte, *supra* note 53, at 676.

225. Field, *supra* note 170, at 645.

226. *Id.* at 653.

227. See, e.g., Gugliuzza, *supra* note 171, at 1457–59 (noting that the Federal Circuit has failed to bring predictability to patent appeals).

228. See, e.g., Dan L. Burk, *Legal and Technical Standards in Digital Rights Management Technology*, 74 *FORDHAM L. REV.* 537 (2005); Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 *DUKE L.J.* 557 (1992); Russell B. Korobkin, *Behavioral Analysis and Legal Form: Rules vs. Standards Revisited*, 79 *OR. L. REV.* 23 (2000); Eric A. Posner, *Standards, Rules, and Social Norms*, 21 *HARV. J.L. & PUB. POL'Y* 101 (1997).

229. See, e.g., Cathryn Johnson, Timothy J. Down & Cecilia L. Ridgeway, *Legitimacy as a Social Process*, 32 *ANN. REV. SOC.* 53 (2006).

230. See, e.g., Richard H. Fallon, Jr., *Legitimacy and the Constitution*, 118 *HARV. L. REV.* 1787, 1795 (2005).

with low levels of sociological legitimacy are expected to rely on codified, rule-like procedures that make diagnoses and treatments more mechanical and transparent, and less reliant on an expert's tacit knowledge.²³¹ In contrast, expert communities that enjoy high levels of sociological legitimacy (and thus higher levels of trust) can rely more heavily on tacit or uncodified contextual knowledge.²³²

Tying rule formalism to legitimacy crises more precisely requires, at a minimum, both carefully documenting public opinion (and the opinion of relevant audiences) about the Federal Circuit's performance and correlating low public-opinion periods with increases in rule formalism. Such an undertaking is beyond the scope of this article. Nevertheless, observations about Federal Circuit behavior are consistent with the hypothesis that it has turned to rule formalism to address legitimacy crises. The Federal Circuit was created to bring consistency and expertise to patent law, which many believed was crippled by widely divergent standards of patentability among circuits and rampant forum-shopping.²³³ Although during the first decade of its existence, members of the patent bar and academic commentators generally agreed that the Federal Circuit was succeeding in bringing uniformity and predictability to patent law,²³⁴ criticism began to mount in the late 1990s.²³⁵ Specifically, several commentators blamed the Federal Circuit for inconsistent, panel-dependent opinions that failed to bring uniformity to patent law.²³⁶ The Federal Circuit's turn to rule formalism closely followed these waves of critiques.²³⁷ This is consistent with the hypothesis that rule formalism was (at least in part) a response to a crisis in sociological

231. See, e.g., Porter, *supra* note 127, at 226–28.

232. *Id.* at 220 (“Arguments within a [strong] community of specialists can be made with a minimum of formality, only a modest concern for rigor, and with frequent recourse to shared, often tacit knowledge.”).

233. See, e.g., Howard T. Markey, *The Federal Circuit and Congressional Intent*, 2 FED. CIR. B.J. 303, 303 (1992) (describing Congress's “express intent” that the Federal Circuit “contribute to increased uniformity and reliability in the field of patent law”); Charles Shifley, *Flawed or Flawless: Twenty Years of the Federal Circuit Court of Appeals*, 2 J. MARSHALL REV. INTELL. PROP. L. 178, 180 (2003) (praising the Federal Circuit for unifying patent law and arguing that internal disagreement within the Federal Circuit “pales in comparison to the conflicts that the patent law had before creation of the Court”).

234. See, e.g., Dreyfus, *supra* note 40, at 74 (“On the whole, the CAFC experiment has worked well for patent law . . .”).

235. See, e.g., Victoria Slind-Flor, *Federal Circuit Judged Flawed*, NAT'L L.J., Aug 3, 1998.

236. *Id.* (“[M]any members of the intellectual property bar . . . accuse the . . . court of unpredictability, claiming that results are often panel-dependent . . .”).

237. See Thomas, *supra* note 7, at 796 (describing the rise of adjudicative rule formalism at the Federal Circuit in the late 1990s).

legitimacy—much like the DSM was thought to have been developed to address psychology’s vulnerability to public and scientific criticism.²³⁸

c) Predicting Internal Relationships: Managing Dissent

Sociological studies of expertise indicate that experts—who have an “engaged commitment” to their area of expertise—are more prone than novices to develop and defend their individual opinions in the face of disagreement.²³⁹ In turn, this suggests that expert judges will be less prone to follow the norm of “consensus” that is theorized to limit dissenting panel opinions.²⁴⁰ The Federal Circuit’s epistemic diversity is likely to make the court even more prone to disagreement among its members than expert communities that share a common technical background that includes many years of education and socialization into a discipline.²⁴¹

Empirical studies of the Federal Circuit have shown that it dissents significantly more often than other circuit courts on issues of patent law—but not on other issues under its jurisdiction.²⁴² And disagreement appears to be growing: Jason Rantanen and Lee Petherbridge have shown that unanimous decision rates have fallen from more than 80% of all opinions to only 60% between 2005 to 2013.²⁴³

These results support the hypothesis that the Federal Circuit is a “weak” expert community with mounting internal divisions in the area of its expertise. Thus, much like the “weak” expert communities studied by sociologists of expertise, the Federal Circuit would be expected to resort to rule formalism as a mechanism to cure or minimize internal divisions. More specifically, if the prediction that rules serve as a tool to manage internal disagreement holds for the Federal Circuit, one would expect, first, that rules

238. Rochelle Dreyfuss has similarly linked the Federal Circuit’s preference for bright line rules leading to more precise decisions to “the court’s focus on public acceptance” and the need (at least in the Federal Circuit’s early years) to “build public support” for the court. Dreyfuss, *supra* note 7, at 816–17.

239. *See supra* Section III.B.

240. *See, e.g.*, Joshua Fischman, *Estimating Preferences of Circuit Judges: A Model of Consensus Voting*, 54 J.L. & ECON. 781 (2011) (including a “cost of dissent” in the judicial utility function to account for the social norm of consensus in three-judge Federal Circuit panels).

241. *See, e.g.*, Porter, *supra* note 127, at 222 (describing the community of high-energy physicists as “remarkably homogeneous, not only in scientific commitments, but even in terms of personal habits, mannerisms, and dress”).

242. Christopher A. Cotropia, *Determining Uniformity Within the Federal Circuit by Measuring Dissent and En Banc Review*, 43 LOY. L.A. L. REV. 801, 815 (2010) (finding that Federal Circuit judges had a 9.28% dissent rate in patent opinions between 1998 and 2009, while other circuits had a significantly lower rate, ranging from 1.14% to 4.56%, and comparable to the Federal Circuit’s dissent rate of 3.51% across all subject areas).

243. Rantanen & Petherbridge, *supra* note 218.

will be more prominent in particularly divisive issues, and second, that overall reliance on rules versus loose standards would increase with mounting disagreement.

The empirical studies carried out to date do not precisely address these two predictions. These studies do not include the pre-1998 period in which the Federal Circuit enjoyed relatively high levels of sociological legitimacy, and do not attempt to measure the prevalence of rules versus standards. Nevertheless, qualitative evidence is consistent with this explanatory framework. The Federal Circuit's tendency to develop rules is particularly salient on issues that have generated a great deal of internal disagreement, such as patentable subject matter or claim construction. On these and other divisive issues, the Federal Circuit has announced rules after hearing a case *en banc*.²⁴⁴ *In fact, the Federal Circuit also bears a larger proportion of its cases en banc than do other Circuit Courts, and routinely requests and considers briefs by amici curiae—a procedure that Ryan Vacca has likened to administrative rule-making.*²⁴⁵ Thus, the *en banc* process by which the Federal Circuit has developed rules in divisive areas also suggests that rule-making is a mechanism to manage internal dissent and seek external legitimacy.

One final, important feature of codification bears emphasizing: jurisdictional control requires a balance between codification and abstraction. Codification can allow for delegation, increase legitimacy, and manage internal dissent, but at the cost of reducing expert autonomy and discretion. And complete codification of expert skills makes expertise irrelevant in the performance of those tasks. In the context of an expert court, extensive reliance on rules can lead external audiences to question the need for expertise. In this context, the Supreme Court's insistence that the Federal Circuit employ flexible standards,²⁴⁶ and its description of its own role in patent law as providing an "outer shell"²⁴⁷ to be filled out by the Federal Circuit's expertise, could be understood as a call for the Federal Circuit to

244. *See, e.g.*, *CLS Bank Int'l v. Alice Corp.*, 717 F.3d 1269 (Fed. Cir. 2013) (*en banc*); *In re Bilski*, 545 F.3d 943 (Fed. Cir. 2008) (*en banc*); *Cyber Corp. v. FAS Techs., Inc.*, 138 F.3d 1448 (Fed. Cir. 1998) (*en banc*).

245. Ryan Vacca, *Acting Like an Administrative Agency: The Federal Circuit En Banc*, 73 MO. L. REV. 733 (2011).

246. *See, e.g.*, Lee, *supra* note 7, at 46 (arguing that the Supreme Court's interventions in patent law call for "holism and contextual engagement," in contrast with the Federal Circuit's preference for inflexible rules).

247. Transcript of Oral Argument at 28, *Alice Corp. v. CLS Bank Int'l*, 134 S. Ct. 2347, 2352 (2014) (No. 13-298) (Breyer, J.).

return to a more active use of its expertise—that is, its contextual, tacit knowledge of patent law and technology.²⁴⁸

* * *

These three aspects of codification present a more complex and nuanced view of how the Federal Circuit uses rules as a mechanism of jurisdictional competition (the external function of rules) and as a means to minimize internal dissent (the internal function of rules). Disentangling whether a particular rule serves to teach, control, legitimate, or cure internal divisions is a complicated task—in part because a rule can serve all of these functions simultaneously, and in part because the change over time from standards to rules (or vice-versa) is difficult to operationalize empirically. Nevertheless, further empirical analysis can serve to more rigorously test these multiple functions of rules in the Federal Circuit and other specialized courts. For example, one could measure whether a court is more likely to prefer bright-line rules over flexible, indeterminate tests in periods of high-judge turnover, in periods with the greatest epistemic diversity among judges, or in periods of crises of negative public opinion. Internal comparisons between the patent and non-patent docket with respect to the court's tendency to rely on rules would also be informative. So would horizontal comparisons with other expert courts, such as those of bankruptcy and tax (although their different position in the judicial hierarchy would complicate data interpretation).

3. *Typecasting*

Typecasting captures the role of *framing* in problem-classification and analysis by expert communities. As Subsection III.A.1 emphasizes, the subjective aspects of a problem enable different communities of experts to frame a problem as best solved by the specific abstract system of their particular community. In the context of jurisdictional competition, *framing* is a tool that allows a community of experts to both defend and expand its jurisdiction. But an expert community's abstract knowledge system also constrains that community's available framings.²⁴⁹ For example, doctors are constrained by their abstract knowledge system to conceptualize a broken

248. See John M. Golden, *The Supreme Court as "Prime Percolator": A Prescription for Appellate Review of Questions in Patent Law*, 56 UCLA L. REV. 657, 662 (2009) (arguing that—given the comparable institutional expertise of the Federal Circuit—the Supreme Court's role in reviewing the Federal Circuit is justified only as combatting "undesirable ossification of legal doctrine").

249. See Steven Shapin, *Here and Everywhere: Sociology of Scientific Knowledge*, 21 ANN. REV. SOC. 289, 292 (1995) (arguing that knowledge acquisition and concept-application is bounded by the "existing structure of knowledge given . . . by their community and within a structure of purposes sustained by their community").

bone as an ailment of the human body, and to look for solutions and analogies in medical textbooks, not in engineering manuals.²⁵⁰ Yet, engineering principles turned out to have clear applications to bone repair.²⁵¹ Expert communities *typecast* a particular problem as similar to other problems already solved within their abstract knowledge system, and thus amenable to the same type of solutions, and may be less likely than generalists to look widely for available solutions.²⁵²

Consider the following examples in the evolution of patent law jurisprudence, which illustrate the rule of typecasting at the Federal Circuit.

a) Professional Typecasting in *Diamond v. Chakrabarty*

In 1980 the Supreme Court issued a decision, *Diamond v. Chakrabarty*,²⁵³ that many believed ushered in the age of biotechnology commercialization.²⁵⁴ In *Chakrabarty*, a divided Court held that living organisms engineered in the laboratory were patent eligible.²⁵⁵ The unpatentability of microorganisms and of living things more broadly, however, had been a tenet of patent law under the “product of nature” doctrine for at least the previous forty years.²⁵⁶

How, then, was this tenet challenged? Peculiar to this story is the fact that Chakrabarty carried out his research at General Electric—a company traditionally focused on physical technologies.²⁵⁷ Challenging this long-standing view required a new analogy that was readily available to those

250. See, e.g., Pedraza-Fariña, *supra* note 91, at 847 (describing resistance from engineers, biologists and funding institutions to an approach to biology that incorporated insights from engineering).

251. See, e.g., Anne Eisenberg, *Replacement Bones, Grown to Order in the Lab*, N.Y. Times, May 27, 2010.

252. See Joseph Fishman, *Creating Around Copyright*, 128 HARV. L. REV. 1333 (2015) (describing research on creativity that shows that problem-solving is constrained by available, familiar solutions to similar problems—or “prior exemplars”).

253. *Diamond v. Chakrabarty*, 447 U.S. 303 (1980).

254. See, e.g., Sally Smith Hughes, *Making Dollars out of DNA: The First Major Patent in Biotechnology and the Commercialization of Molecular Biology, 1974–1980*, 92 ISIS 541, 569 (2001) (showing that *Chakrabarty* was a “critical ruling for commercial biotechnology”).

255. *Chakrabarty*, 447 U.S. 303

256. The “product of nature” doctrine, which was long thought to block the patentability of living things, dates back to 1889 when the U.S. Commissioner of Patents rejected the application for a patent on a fiber found in a needle of a pine tree. Ex Parte Latimer, 12 Mar. 1889, C.D., 46 O.G. 1638, U.S. Patent Office, *Decisions of the Commissioner of Patents and of the United States Courts in Patent Cases . . . 1889* (Washington, D.C., 1890), 1230127.

257. Daniel J. Kevles, *Ananda Chakrabarty Wins a Patent: Biotechnology, Law, and Society, 1972–80*, 25 HIST. STUD. PHYS. BIOL. SCI. 111, 114 (1994).

working with mechanical inventions and within a different patent culture.²⁵⁸ To most biological scientists and their attorneys, microorganisms—even those manipulated in the laboratory—were not different in kind from products of nature, and were thus unpatentable.²⁵⁹ But to engineers and their patent attorneys, microorganisms manipulated in the laboratory could be analogized to physical products made of different parts.²⁶⁰ Once this new analogy was articulated, it became possible—and even a matter of simple legal logic—to think of living organisms not as natural products, but as items manufactured out of chemical subunits.²⁶¹

This example does not directly or uniquely involve the Federal Circuit as an expert community. It does, however, illustrate that different expert communities (in this case patent attorneys specializing in biotechnology versus those specializing in mechanical products) are bounded by their most readily available framing of a problem. The next example directly concerns the Federal Circuit and the PTO.

b) Technical Typecasting in *Myriad Genetics*

Following the Supreme Court's decision in *Diamond v. Chakrabarty*, the PTO began granting patents to isolated DNA sequences²⁶²—analogizing DNA sequences that had been extracted from an organism to purified chemical compounds, which had long enjoyed patent protection.²⁶³ But DNA, and specifically DNA sequences within a gene, could also be analogized to an information carrier whose main role is to hold and transmit information, rather than participate in chemical reactions.²⁶⁴ Neither the PTO nor the Federal Circuit appears to have given much consideration to this

258. According to *Chakrabarty*, “companies like major drug firms, long accustomed to the product-of-nature barrier to patents, would not have filed a patent application on his new bugs.” 447 U.S. at 117.

259. *Id.*

260. *Id.*

261. *See, e.g.*, *In re Application of Bergy*, 596 F.2d 952, 974–75 (“The nature and commercial uses of biologically pure cultures of microorganisms . . . are analogous in practical use to inanimate chemical compositions . . . used in the chemical industry.”).

262. *See, e.g.*, U.S. Patent 4,370,417 (filed Jan. 25, 1983) (claiming DNA sequence for plasminogen activator protein); U.S. Patent 4,703,008 (filed Oct. 27, 1987) (claiming DNA sequence for erythropoietin).

263. *See* Utility Examination Guidelines, 66 Fed. Reg. 1092, 1093 (Jan. 5, 2001) (“Like other chemical compounds, DNA molecules are eligible for patents when isolated from their natural state and purified or when synthesized in a laboratory from chemical starting materials.”).

264. *See, e.g.*, Rebecca S. Eisenberg, *How Can You Patent Genes?* 2 AM. J. BIOETHICS 3, 9 (2002) (“The DNA molecule itself may be thought of as a tangible storage medium for information about the structure of proteins.”).

distinction.²⁶⁵ Rather, it was the Southern District of New York in its *Myriad* decision that engaged in a discussion of the implications of the information carrier analogy for the patentability of genes.²⁶⁶ The district court concluded that the DNA-as-information-carrier analogy rendered isolated genes unpatentable products of nature.²⁶⁷

The Federal Circuit considered the *Myriad* case twice—once on appeal from the Southern District of New York and again on remand from the Supreme Court, which instructed the Federal Circuit to reconsider the case in light of its decision in *Mayo v. Prometheus*.²⁶⁸ As emphasized in Part IV.B.1.c), an application of the reasoning in *Mayo* to *Myriad* could have led the CAFC to focus on the informational content of DNA. Nevertheless, Judge Lourie focused on the molecular structure of genomic DNA, framing DNA as a molecule with a “distinctive chemical structure and identity from those found in nature,” rather than an information carrier.²⁶⁹ Under this framing, Judge Lourie concluded that “*Mayo* does not control the question of patent eligibility.”²⁷⁰ Isolated DNA is not a “product of nature” because it “exists in a distinctive chemical form—as distinctive chemical molecules—from DNAs in the human body.”²⁷¹ Judge Lourie holds a Ph.D. in chemistry and it is plausible that his views in this case are shaped and filtered through his previous technical training.²⁷² Arti Rai has similarly hypothesized that Judge Lourie’s obviousness analysis of DNA-based inventions was influenced by his technical background in chemistry.²⁷³

265. See, e.g., *Ass’n for Molecular Pathology*, 689 F.3d 1303 (Fed. Cir. 2012) (“It is undisputed that Myriad’s claimed isolated DNAs exist in a distinctive chemical form—as distinctive chemical molecules—from DNAs in the human body.”); Utility Examination Guidelines, 66 Fed. Reg. at 1093 (“A purified DNA *molecule* isolated from its natural environment, on the other hand, is a chemical compound and is patentable if all the statutory requirements are met.”).

266. *Ass’n for Molecular Pathology v. U.S. Patent & Trademark Office*, 702 F. Supp. 2d 181, 228–29 (S.D.N.Y. 2010), *aff’d in part, rev’d in part*, 653 F.3d 1329 (Fed. Cir. 2011), *aff’d in part, rev’d in part*, 689 F.3d 1303 (Fed. Cir. 2012), *aff’d in part, rev’d in part sub nom. Ass’n for Molecular Pathology v. Myriad Genetics, Inc.*, 133 S. Ct. 2107 (2013).

267. 702 F. Supp. 2d at 229 (concluding that the “defining characteristic” of DNA was its role as a “physical embodiment of information,” and that “the preservation of this defining characteristic of DNA in its native and isolated forms mandates the conclusion that the challenged composition claims are directed to unpatentable products of nature”)

268. *Mayo*, 132 S. Ct. 289.

269. *Ass’n for Molecular Pathology*, 689 F.3d at 1328.

270. *Id.*

271. *Id.*

272. *Patent Gold Rush*, *supra* note 8, at 218 n.64 (noting that Judge Lourie has a Ph.D. in Chemistry).

273. *Id.*

These two examples also represent two types of typecasting that can operate at the level of the Federal Circuit. The first, professional typecasting, refers to the possibility that the prior professional embeddedness of a judge in a particular community (for example, the patent law community) may influence that judge's interpretation of a problem.²⁷⁴ The second, technical typecasting, refers to the possibility that the framing of a problem is influenced by previous technical training in a particular field. Professional and technical typecasting provide a finer-grade description of two types of "tunnel vision" that can afflict an expert court.

Typecasting is undoubtedly a helpful tool in allowing expert communities to sharply define an issue as analogous to other issues for which an expert community has already found answers, making it amenable to solutions using well-established methods. But typecasting has a clear downside, in particular in the context of a centralized court. As the struggle between psychiatrists, psychologists, and the clergy to treat alcoholism illustrates, different professions compete in the marketplace for the demand of their services by framing particular problems as best solved by the tools of their expertise. That competition among alternative frames is lacking in patent law. I analyze the normative implications of typecasting in the last Part of this Article.

4. *Inability to self-coordinate across multiple expert areas*

This last feature of expert communities combines two insights from the sociology of expertise as embodied intuition and from the sociology of the professions. First, an expert's "engaged commitment"²⁷⁵ with her area of expertise suggests that issues related to that area of expertise are particularly personally salient to experts relative to the general public, and relative to issues in other fields of expertise.²⁷⁶ In turn, this propensity to care more about, and thus focus more on, an expert's field of study makes it less likely for experts in one area to pay adequate attention to problems and solutions within other areas of expertise. When coordination with other expert areas requires trade-offs—as is the case with patent law and antitrust where, for example, protecting consumers from anti-competitive settlements or

274. *Id.*

275. *See supra* Section II.B.

276. This is consistent with Judge Richard Posner's observations that a specialized judiciary would "attract persons of somewhat different abilities . . . who are more deeply interested in particular subjects and less interested in running everything." RICHARD A. POSNER, *THE FEDERAL COURTS* 250 (1996).

practices may require constraining patent entitlements²⁷⁷—a community with expertise in one area may place inadequate weight on the competing interests of other expert communities.

Second, competition to fully occupy an expert space (i.e., to attain full control over a jurisdiction) often prevents spontaneous, sustained cooperation among expert communities with different abstract knowledge bases.²⁷⁸ When such cooperation is required—as is, for example, in “wicked problems”²⁷⁹ that require action across multiple areas of expertise—it will be difficult for expert communities to self-coordinate across multiple expert areas. Instead, coordination will likely require external structuring or incentives.

Coordination challenges at the Federal Circuit have given rise to a different type of tunnel vision. Generally, innovation policy has turned out to be a “wicked problem.” Incentivizing innovation was one of the key driving forces behind the creation of the Federal Circuit, but patent policy is but a single piece in the mosaic of policies designed to encourage innovation. Thus, knowledge required for fashioning innovation policy that is attentive to the welfare-maximizing balance between patent protection and market competition resides in multiple government institutions.²⁸⁰ And, as Stuart Benjamin and Arti Rai have argued, courts and agencies that regulate innovation are often unaware of each other’s solutions to similar problems.²⁸¹

At a fundamental level, coordination challenges concern the organization of knowledge in isolated communities or isolated institutions.²⁸² Trans-

277. See, e.g., Mark A. Lemley, *Industry-Specific Antitrust Policy for Innovation*, 2011 COLUM. BUS. L. REV. 637, 640 (2011) (“IP and competition frequently work at cross purposes toward the goal of innovation.”).

278. See Pedraza-Fariña, *supra* note 91, at 845 (arguing that knowledge exchange among scientific communities is significantly impaired by individual communities’ “resistance to ‘outside’ tools and interpretive frameworks”). See also Gary Becker & Kevin Murphy, *The Division of Labor, Coordination Costs, and Knowledge*, in *HUMAN CAPITAL: A THEORETICAL AND EMPIRICAL ANALYSIS WITH SPECIAL REFERENCE TO EDUCATION* 299, 308–09 (Gary S. Becker, ed., 1994) (describing coordination problems arising from knowledge specialization).

279. See, Horst Rittel & Melvin Webber, *Dilemmas in a General Theory of Planning*, 4 POL. SCI. 155, (1973) (arguing that the specialization of labor and expertise has failed to solve “wicked problems,” such as poverty, crime, and public education, whose interconnectivity and complexity requires a coordinated approach).

280. See, e.g., Evan Selinger & Thomas Seager, *The Incompatibility of Industrial Age Expertise and Sustainability Science*, in *EXPERTISE: PHILOSOPHICAL REFLECTIONS* 99, 106 (Evan Selinger, ed., 2011).

281. Stuart Minor Benjamin & Arti K. Rai, *Fixing Innovation Policy: A Structural Perspective*, 77 GEO. WASH. L. REV. 1 (2008).

282. *Id.* at 19.

institutional knowledge is required for developing innovation policy but access to such knowledge is “significantly handicapped by the degree to which [it] resides in increasingly narrow specializations [or institutions].”²⁸³ This represents a second type of tunnel vision, distinct from typecasting.

But coordination difficulties are not only about lack of awareness of solutions, but also about preferences for, or emotional attachments to, a particular approach to a problem. In this sense, the “inability to coordinate” and “typecasting” features of expert communities are linked: failure to coordinate may be due to a refusal to accept an alternative framing as valid, or to accord it sufficient weight.

Take, for example, the tension between competition law and patent law. In her 1989 analysis of the Federal Circuit’s performance in the five years following its creation, Rochelle Dreyfuss pointed out a coordination problem that persists to this day: “If the CAFC is told to encourage invention, but is permitted to see only a small part of the matrix into which patent cases fit [i.e., only patent law] . . . it will undervalue the interest of competitors because it will not have the occasion to consider the role that vigorous competition plays in encouraging invention.”²⁸⁴ Dreyfuss’s analysis implied that expanding the Federal Circuit’s jurisdiction to include antitrust cases would correct this imbalance.²⁸⁵ More recently, Paul Gugliuzza took a similar stance in suggesting that replacing some of the Federal Circuit’s non-patent docket with commercial disputes may improve the Federal Circuit’s understanding of the place of patent law within the broader array of policies designed to incentivize innovation.²⁸⁶

Unlike epistemic control and codification—features which, taken alone, do not have clear normative implications for the design of the Federal Circuit—typecasting and inability to self-coordinate have normatively undesirable consequences in the context of a centralized court. I explore the predictions and normative implications of these two features in the last Part of this article.

* * *

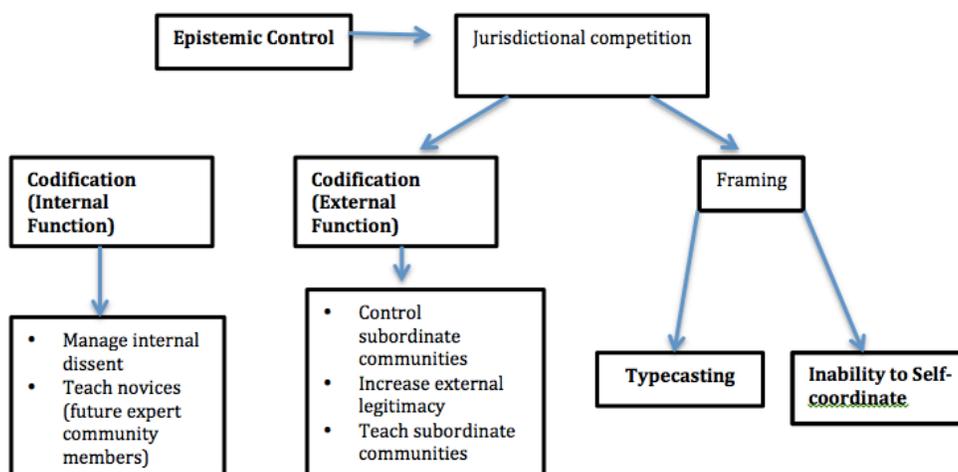
283. Selinger & Seager, *supra* note 280, at 106.

284. Dreyfuss, *supra* note 40, at 54.

285. *Id.*

286. Paul R. Gugliuzza, *Rethinking Federal Circuit Jurisdiction*, 100 GEO. L.J. 1437, 1498 (2012).

Figure 1 below diagrams the relationships among the four features of the typology:



C. ALTERNATIVE EXPLANATIONS OF FEDERAL CIRCUIT BEHAVIOR

Commentators have advanced several explanations for some of the puzzling features of the Federal Circuit described in the previous sections: (1) centralization theory, (2) information-costs theory, and (3) capture and tunnel vision. Nevertheless, none of these theories can fully account for Federal Circuit behavior. This next section engages with these explanations and contends that the expert-community model outlined here both complements these explanatory frameworks and is necessary to fully understand the Federal Circuit.

1. *Centralization Theory*

Several scholars have attributed the Federal Circuit's shortcomings to its central position as the single court for patent appeals.²⁸⁷ For example, Craig Nard and John Duffy argue that the Federal Circuit has achieved uniformity at the expense of quality.²⁸⁸ According to Nard and Duffy, the Federal

287. See, e.g., Nard & Duffy, *supra* note 47, at 1622; Hon. Wood, *supra* note 6; Dreyfuss, *supra* note 12.

288. *Id.* at 1620 (arguing that Federal Circuit precedent does not “adequately reflect[] current knowledge regarding the beneficial functions of the patent system in generating technological innovation, the potential problems of patent rights in foreclosing legitimate competition, and the need for predictable rules capable of curtailing litigation costs”).

Circuit's mandate to achieve uniformity in patent law has resulted in decisions that are divorced from the needs of the very communities whose innovation patent law is supposed to incentivize.²⁸⁹ Nard and Duffy argue that a centralized appeals structure facilitates not only uniformity but also isolation and lack of experimentation with novel approaches to patent law.²⁹⁰ Their proposed solution is to reconfigure the centralized structure of patent appeals to re-introduce a measure of competition and diversity that would inject more incremental innovation and flexibility into the patent system.²⁹¹ Similarly, in a recent keynote address, Chief Judge Diane Wood argued for the reintroduction of "the same kind of marketplace of ideas [in patent law] at the court of appeals level that we have for almost every other kind of claim"²⁹² by allowing parties to file their case either in the Federal Circuit or in the regional circuit where their claim was first filed.²⁹³ Rochelle Dreyfuss has also argued that lack of percolation is a significant problem that has had a negative impact on the quality of Federal Circuit patent decisions.²⁹⁴ Dreyfuss has proposed according greater deference to PTO decision-making, and to select District Court judges that participate in the Patent Cases Pilot Program.²⁹⁵

Nevertheless, taken alone, centralization and the drive for uniformity cannot explain many of the features of the Federal Circuit described above.²⁹⁶ First, the Federal Circuit often disregards its own "rules," a fact that has led many district court judges to express their frustration with the appellate court.²⁹⁷ Strict adherence to the uniformity principle would not predict such a malleable interpretation of its own precedent. In addition, the Federal Circuit

289. *Id.*

290. *Id.* at 1649.

291. *Id.* at 1623, 1625 (proposing that "both the Federal Circuit and United States Court of Appeals for the District of Columbia Circuit (D.C. Circuit) . . . have jurisdiction over appeals from the PTO").

292. Hon. Wood, *supra* note 6, at 9.

293. *Id.*

294. Dreyfuss, *supra* note 12, at 524–26.

295. *Id.* at 532–36.

296. Duffy and Nard do not advance centralization as a comprehensive explanatory framework for Federal Circuit behavior. *Supra* note 47.

297. Inconsistent rule application is most salient in issues of claim construction and choice of law rules. *See, e.g., A Panel Discussion: Claim Construction from the Perspective of the District Judge*, 54 CASE W. RES. L. REV. 667, 678 (remarking that the Federal Circuit applies its rules inconsistently to claim construction, making it difficult for district court judges to avoid reversal); Ted Field, *Improving the Federal Circuit's Approach to Choice of Law for Procedural Matters in Patent Cases*, 16 GEO. MASON L. REV. 643, 644–53 (2009) (explaining the Federal Circuit's inconsistent application of choice-of-law rules).

has an unusually high rate of dissent for an appellate tribunal²⁹⁸—a fact that is not easily reconciled with a court for which uniformity is of paramount importance.²⁹⁹ These features are best explained as a consequence of institutionalized expert decision-making.³⁰⁰ Second, it is unclear how centralization and the uniformity principle explain the Federal Circuit’s relationship with other expert bodies, such as the PTO and the ITC.³⁰¹ The low level of deference accorded to those institutions has arguably done much to undermine the predictability and uniformity of patent law.³⁰² But there is nothing inherent in the concept of centralization that predicts this low-level of deference to institutions that have a considerable level of expertise in patent law issues.³⁰³

2. *Information-Costs Theory*

Peter Lee has advanced an “information-cost theory” of the Federal Circuit that explains adjudicative rule formalism as a heuristic to manage the cognitive burdens and technological anxieties of generalist district court judges.³⁰⁴ Under this account, the Federal Circuit prefers rigid rules to flexible standards because rules diminish the need for lay judges to engage deeply with complex technologies.³⁰⁵ Similarly, John Thomas argues that simple rules “might be seen as providing a well-meaning judiciary with a thread through the labyrinth [of complex patent law].”³⁰⁶ And Rochelle Dreyfuss has

298. See, e.g., Rantanen & Petherbridge, *supra* note 218.

299. Nard and Duffy recognize that dissents can be a source of divergent opinions, but argue that they are insufficient to overcome the pull of circuit precedent, and thus not as efficient in creating legal innovation as a decentralized system of appellate courts. Nard & Duffy, *supra* note 47, at 1646.

300. See *supra* Section III.A. This article does not argue that centralization and the drive for uniformity do not play a role in explaining Federal Circuit behavior. Rather, it argues that a conceptualization of the Federal Circuit as an institutionalized community of experts helps explain a host of additional puzzling behaviors and provides an additional lens by which to understand, judge, and design the institutions in charge of administering patent law.

301. See *supra* Subsection IV.C.1

302. See *supra* Subsection IV.C.2

303. Centralization can predict low levels of deference to expert agencies if we assume that centralized courts will seek to consolidate top-down decision-making power and seek final decision-making authority. But it is equally possible that a centralized court designed to increase predictability and uniformity in patent law would choose to defer to expert agencies to set coherent and predictable patent policy. Conceptualizing a centralized court as an expert court that competes with other expert communities for epistemic control over patent law provides a more direct link between the explanatory framework (expert community theory) and the observed behavior (lack of deference to other expert communities).

304. Lee, *supra* note 7, at 25.

305. *Id.* at 9.

306. Thomas, *supra* note 7, at 795.

explained the Federal Circuit's resort to bright line rules as an aid for "trial judges [who] do not have the expertise enjoyed by specialized jurists . . . [to] treat like cases alike, even if some of the complexities in the technological parts of the cases are obscure."³⁰⁷ As elaborated in Part V, an expert decision-making model places the actions of the Federal Circuit within a broader framework. Consistent with Lee's information-cost theory, expert communities resort to rules to codify and simplify expert knowledge for external, lay consumption. But rules also serve two additional purposes. They act as gatekeepers of an expert community's jurisdictional power, by implicitly stating that outsiders do not possess the requisite know-how to correctly engage with a particular subject matter, in this case, technology policy through patent law.³⁰⁸ Rules can also enhance the legitimacy of weak expert communities. Because rules arguably reduce the influence of subjective factors in decision-making, they serve to manage internal dissent and to increase external support.³⁰⁹

3. *Capture and Tunnel Vision*

A final explanatory framework used to describe the Federal Circuit's behavior relies on its status as a "specialized" court. Indeed, the dangers of a specialized judiciary appeared prominently in debates leading to the creation of the Federal Circuit.³¹⁰ But Congress focused almost exclusively on two features of a specialized judiciary that are thought to negatively influence decisional content—capture and tunnel vision.³¹¹ Capture describes the

307. Dreyfuss, *supra* note 12, at 512.

308. *See supra* Part IV.B.1.b).

309. *See supra* Part IV.B.2.c).

310. *See, e.g.,* H.R. Rep. No. 97-312, at 31 (1981) ("Several witnesses before the Committee expressed fears that the Court of Appeals for the Federal Circuit would be unduly specialized or would soon be captured by specialized interests. This provision should reduce these fears by ensuring that all the judges sit on a representative sampling of all the cases heard.")

311. *See, e.g.,* *Hearing on S. 21 and S. 537 Before the Subcomm. on Courts of the S. Comm. on the Judiciary*, 97th Cong. 211 (1981) ("The quality of decision-making would suffer as specialized judges become subject to 'tunnel vision' seeing the cases in a narrow perspective without the insights stemming from broad exposure to legal problems in a variety of fields."); S. Rep. No. 97-275, at 6 (1981) ("[T]he subject matter of the new court will be sufficiently mixed to prevent any special interest from dominating it."); Dreyfuss, *supra* note 40, at 3 (noting that critics of specialization argue it "will produce a court with tunnel vision, with judges who are overly sympathetic to the policies furthered by the law that they administer or who are susceptible to 'capture' by the bar that regularly practices before them"); William Landes & Richard Posner, *An Empirical Analysis of the Patent Court*, 71 U. CHI. L. REV. 111, 111–112 (2003) (positing that a specialist court may be captured because "interest groups that had a stake in patent policy would be bound to play a larger role in the appointment of judges of such court than they would in the case of the generalist federal courts."); Simon Rifkind, *A*

external influence of interest groups on the policies and decisions of a particular institution.³¹² It is not, however, uniquely linked to specialization. Both centralization and specialization can facilitate capture by special interest groups. The former does so by making it easier to coordinate and focus lobbying activities on a small number of judges; the latter because specialized judges are likely to be part of the same professional network with repeat industry players.

Congress was also concerned with what it has termed the “undue specialization” of the Federal Circuit.³¹³ Concerns about undue specialization relate to internal characteristics of specialized bodies that are thought to negatively impact decision-making. Thus, Congress often referred to the “narrowness,” “technical focus” and “tunnel vision” that may arise from specialization.³¹⁴ Counteracting such narrowness required exposing judges to cases from a variety of fields,³¹⁵ and ensuring the court was not simply staffed by “patent lawyers.”³¹⁶ But tunnel vision is an ambiguous concept. It hides

Special Court for Patent Litigation?, The Danger of a Specialized Judiciary, 35 A.B.A. J. 425, 425 (1951) (“Once you segregate the patent law from the natural environment in which it now has its being, you contract the area of its exposure to the self-correcting forces of the law.”).

312. William J. Novak, *A Revisionist History of Regulatory Capture*, in PREVENTING REGULATORY CAPTURE: SPECIAL INTEREST INFLUENCE AND HOW TO LIMIT IT 25 (Daniel Carpenter & David Moss eds., 2013).

313. *See, e.g.*, H.R. Rep. No. 97-312, at 31 (“Several witnesses before the Committee expressed fears that the Court of Appeals for the Federal Circuit would be *unduly specialized* . . .”) (emphasis added); *id.* at 50 (“This amendment . . . does, however, clearly send a message to the President that he should avoid undue specialization [sic] on both courts.”; S. Rep. No. 97-275, at 6 (rejecting a proposal to expand the Court of Customs and Patent appeals as “inconsistent with the imperative of avoiding undue specialization within the Federal Judicial system”).

314. *See, e.g.*, Report of Committee of the Section of Patent, Trademark, and Copyright Law to the Section of Patent Law of the ABA, at 548 (“The proposed method of making up the Court will obviate the principal objection which exists to the creation of a court of patent appeals . . . which is, that a permanent court consisting of judges appointed for life and occupied in the sole work of deciding patent cases would be liable to grow narrow and technical in its views and procedure.”).

315. *See, e.g.*, H.R. Rep. No. 97-312, at 19 (“By combining the jurisdiction of the two existing courts along with certain limited grants of new jurisdiction, the bill creates anew intermediate appellate court markedly less specialized than either of its predecessors and provides the judges of the new court with a breadth of jurisdiction that rivals in its variety that of the regional courts of appeals.”).

316. *See, e.g.*, Report of Committee of the Section of Patent, Trademark, and Copyright Law to the Section of Patent Law of the ABA, *supra* note 315, at 548–49 (“Under the proposed plan the judges would come to the court of patent appeals trained for their work by experience on the bench in the field of general jurisprudence. It would give us a court of judges, and not of mere patent lawyers.”).

multiple potential mechanisms for influencing the content of judicial decisions. Consider, for example, the following five:

1. The professional biases of patent lawyers towards regarding patents as valid will cause judges with a background in patent practice to favor patent holders,³¹⁷ analyzed by this article as a form of *professional typecasting*.
2. The technical background of judges will influence how they evaluate the worth, and thus patentability, of particular inventions, analyzed by this article as *technical typecasting*.
3. Judges who are experts in patent law are unable to fully grasp and consider the impact of their decisions on other fields of law, notably competition law, analyzed by this article as an *inability to self-coordinate*.
4. Judges get so used to a particular way of approaching problems within their expertise they no longer question, or are willing to question, the validity of their foundational assumptions, analyzed by this article as a consequence of *epistemic monopoly and epistemic autonomy*.
5. Expert judges will no longer be understood by non-experts—and thus their work will not be transparent and easily accessible to lay people. Thus raising issues of public *trust* and the *legitimacy* of expert decision-making.³¹⁸

Whether and how “tunnel vision” should be corrected depends on understanding the specific mechanisms through which it influences decision-making.

Capture and tunnel vision are also insufficient to explain the specific features of Federal Circuit jurisprudence described above. First, capture is overinclusive, as it describes behavior linked both to centralization and specialization. Most importantly, although the Federal Circuit has long been viewed as a pro-patent court, many of its decisions have limited the scope of

317. See, e.g., LAWRENCE BAUM, SPECIALIZING THE COURTS 36 (2011) (arguing that people who work in patent law are likely to have “a narrower range of opinions about the issues in their field than does the general public or political and social elites as a whole”).

318. It is to this aspect of “tunnel vision” that Judge Wood appears to be referring to when remarking: “Law, in the final analysis, governs society. It should not be an arcane preserve for specialists, who never emerge to explain, even to their clients, what the rules are or why one side or the other prevailed.” Wood, *supra* note 6, at 7.

patent grants, thus undercutting the explanatory power of capture theory.³¹⁹ Further, neither capture nor tunnel vision can fully explain the interactions between the Federal Circuit and other judicial and administrative bodies, nor its preference for rules over standards.³²⁰

The next section analyzes normative proposals to redesign the Federal Circuit in light of this Article's model of expert decision-making.

V. NORMATIVE IMPLICATIONS: EVALUATING PROPOSALS FOR FEDERAL CIRCUIT REFORM

Two features of expert communities—epistemic control and codification—taken alone do not have clear normative implications for the design of the Federal Circuit. Whether or not it is normatively desirable to allow the Federal Circuit to take a non-deferential stance towards district courts and other agencies depends in large part on whether the Federal Circuit has not only attributed expertise in patent law, but also has real, substantive expertise that allows it to reach a better solution—given the goals of patent law—than other institutional players.

“Expertise” could be defined simply as an attribution. This is the view taken by some sociologists.³²¹ Under this relational view of expertise, being an expert means that others attribute the quality of expertise to a particular community or individual which also views itself or herself as expert.³²² In contrast, a substantive view of expertise considers expertise a real, substantive attribute that can lead to “better results” in solving particular social problems.³²³ To be an expert means to “know what you are talking

319. *See, e.g.*, Dreyfuss, *supra* note 40, at 28 (noting that in the five years following its creation the Federal Circuit adopted rules “which favor technology users,” such as stringent reviews of practice before the PTO and a restrictive interpretation of the doctrine of equivalents).

320. The analysis advanced here is complementary to the model of specialized courts proposed by political scientist Lawrence Baum as it focuses not on individual actors, as Baum does, but on the aggregate behavior of expert institutions and groups of experts. Importantly, key insights emerge when we study the activities, work, and discourse of communities of experts, rather than individuals themselves. For example, Baum's approach does not attempt to explain rule-formalism or how expert communities (such as the PTO, ITC, and the Federal Circuit) relate to each other. *See* Baum, *supra* note 317.

321. *See, e.g.*, Collins & Evans, *supra* note 135, at 2 (describing the “sociology of the acquisition of expert status” as showing that “coming to be called an expert may have little to do with the possession of real and substantive expertise”).

322. *Id.* at 2 (describing relational approaches to expertise as the “assignment of a label”).

323. *Id.* at 2–3 (describing Collins and Evans view of expertise as “the real and substantive possession of groups of experts”).

about”³²⁴ more than non-experts by some external, objective measure—not simply by convincing others.³²⁵ The typology of expert decision-making described in this article sits comfortably with either view of expertise—it applies to communities that perceive themselves (and are perceived) as expert whether or not that expertise is substantively justified. Indeed, attributional disputes over who has expertise and thus should have control over a particular jurisdiction are at the heart of the jurisdictional competition view of expertise. But jurisdictional disputes are often resolved when one community gains legitimacy over another in the eyes of the public. Under a substantive view of expertise, public legitimacy is based, at least in part, on substantive claims of efficiency of results—which are capable of external, objective verification.

This Article takes the substantive view of expertise, that is, that experts can “know what they are talking about” more than non-experts. Indeed, a view of expertise as purely relational would defeat any efforts to improve the quality of expert decision-making. Ultimately, if expertise is a substantive quality, it is not enough to label a community as expert simply because itself and others perceive it as such: there must be some external, objective measurement of expertise. That external measure can also help guide how to best construct expert decision-making bodies. But this Article brackets an analysis of the content of expertise, and concerns itself with modeling the behavior of those communities that are considered expert under the broader relational view. This is a crucial first step in a broader project, not in the least because engaging substantively with how to foster the right type of expertise will require a field-by-field analysis (that is, an expert tax court may look quite different from an expert patent court).

Crucially, this broader descriptive project has normative implications insofar as it identifies characteristics of all expert communities that are likely to be normatively undesirable, suggests avenues for addressing them, and points out why some of the current solutions to the specialization problem are likely to fail. More specifically, the next Part analyzes how the features of typecasting and the inability to self-coordinate have normatively undesirable consequences in the context of centralized courts. Taking them into account can both help evaluate and design proposals for Federal Circuit reform. The model presented in this Article calls into question the effectiveness of proposals that seek to improve Federal Circuit decision-making by expanding its docket to include competition cases, or by allowing a second, or third,

324. *Id.*

325. *Id.*

appellate court to hear patent cases. This Part closes by outlining an alternative proposal for Federal Circuit design: the use of external advisory panels as a solution to the problems of typecasting and coordination.

A. THE DANGERS OF TYPECASTING IN A CENTRALIZED EXPERT COURT

Typecasting can act as a heuristic that formulates what may otherwise be an intractable problem into a solvable question. As a framing device, typecasting is an important tool in efficient problem-solving within an expert community.

When expert communities compete for the demand of their services in the professional world, they effectively pit their framing devices against each other as the most effective means to solve particular problems—seeking to gain legitimacy in the eyes of relevant audiences: consumers of their services and lawmakers with the power to alter rules in their favor.³²⁶ Thus, the market for services effectively tests expert communities' claims that their approach leads to the best results.³²⁷

But this type of weeding-out mechanism doesn't function, or is severely impaired, in a centralized expert court embedded in a court hierarchy for two key reasons. The first is the absence of competition between alternative frames through competition among *peer* expert communities. Despite the epistemic diversity of the Federal Circuit, an important goal of any community of experts, including the courts of appeals,³²⁸ is to reach consensus on their approach to a particular problem. And as John Duffy and Craig Nard have pointed out, the Federal Circuit lacks any peer institutions that can be effective competitors: although both the PTO and the ITC have expertise in patent law they are subordinate, rather than peer, expert communities.³²⁹ The second is the pull of precedent—once consensus is reached and announced in a judicial opinion, framing devices become sticky.³³⁰ Take, for example, the Federal Circuit's treatment of DNA. Framing DNA by reference to its chemical structure has permeated the Federal Circuit's analysis of DNA first in the obviousness inquiry and later when considering patentable subject matter.³³¹ Thus, in the context of a centralized court, typecasting is likely to lead to lower quality decisions, in particular by

326. See *supra* Part IV.A

327. See, e.g., Abbott *supra* note 108, at 140 (arguing that “jurisdictional contests are often decided by client choice”).

328. See, e.g., Fischman, *supra* note 240.

329. Nard & Duffy, *supra* note 47, at 1637.

330. See *id.* at 1630.

331. See *supra* note 276.

preventing alternative framings (and thus solutions) of a problem from being fully explored.

One solution proposed by John Duffy and Craig Nard, is to decentralize judicial decision-making in patent law by allowing two to four circuit courts, including the Federal Circuit, to hear patent cases.³³² This would allow a measure of competition between alternative frames, and thus diminish the problem of typecasting. It is unclear, however, whether the effect of different judicial methodologies or framings on innovation can be efficiently assessed, given the national and often international nature of innovative activity. In other words, it would be a very difficult task to attribute a specific, differential real-world effect on innovation to differences in judicial approaches. Hon. Judge Diane Wood also made a proposal to decentralize patent appeals.³³³ Hon. Judge Wood's proposal, however, would allow appellants to choose to file their appeal either with the Federal Circuit or the regional circuit court.³³⁴ Allowing the parties to choose the forum may provide a quality-control mechanism similar to the market for services in the professional realm. This mechanism, however, is imperfect since parties with weak claims may choose a forum not because of its efficiency or accuracy in rendering patent decisions, but simply because the forum is perceived as patent-friendly, or vice-versa.³³⁵

Decentralization may also not cure typecasting if the composition of the courts is quite similar (e.g., if the courts are partially staffed by former patent attorneys and academics, or those with a technical background, or if the courts develop expertise in patent law through prolonged, concentrated exposure to patent cases)—as these courts may ultimately employ similar frames when reaching patent decisions. On the other hand, if decentralization is achieved by funneling some patent cases to generalist judges, it would come at the cost of losing expertise. To the extent that a court with real, substantive expertise in patent law would reach better decisions than a court of generalist judges, at least most of the time, this is a real concern.

332. Nard & Duffy, *supra* note 89, at 1629–37.

333. Hon. Wood, *supra* note 6, at 9.

334. *Id.*

335. Hon. Diane Wood proposes that, in cases where both parties are aggrieved, jurisdiction among the Federal Circuit and the Circuit where the case was originally filed could be decided by the Judicial Panel on Multidistrict Litigation (“JPML”) to avoid forum-shopping. *Id.* See also Rochelle Cooper Dreyfuss, *Abolishing Exclusive Jurisdiction in the Federal Circuit: A Response to Judge Wood*, 13 CHI.-KENT J. INTELL. PROP. 327, 342 (2014) (discussing selection effects).

A second solution may be to increase the diversity of relevant technical and professional backgrounds in the court with the goal of representing key innovation sectors and approaches to innovation policy. One option may be to appoint more judges with technical expertise in software design or professional backgrounds in antitrust law. But appointing judges with particular technical expertise is likely ill-advised.³³⁶ The structural constraints of a court of appeals regarding the number of judges, currently twelve, make it impossible to appoint judges with expertise in every single area of technology that comes before the court. And even if such constraints did not exist, or if they could be circumvented (for example, by a system of rotating technical judges with expertise in particular technology areas), the rapidly evolving nature of scientific research makes this proposal impracticable. Scientific fields are not static; in fact, new fields of scientific inquiry often redraw the boundaries between technical specialties, making it hard to match judicial technical expertise with case background.³³⁷ And expertise in a scientific field, and in particular in fast-moving fields, is quickly eroded when a judge ceases to be embedded in the relevant scientific community.³³⁸

B. OVERCOMING COORDINATION DIFFICULTIES

Incentivizing innovation requires not only coordinating across different institutions such as the PTO, ITC, Federal Circuit, Food and Drug Administration, and the Federal Trade Commission, but also balancing often-opposing goals within single institutions. In patent law, the most important tension is between patent protection and free competition. For this reason, proposals to re-balance the Federal Circuit's caseload often involve adding to it antitrust cases.³³⁹ Nevertheless, the analysis offered here gives reasons to be skeptical that simply broadening the Federal Circuit's docket to include commercial disputes would lead to better coordination between patent and antitrust. In particular, because the Federal Circuit already views itself (and is viewed by outside observers) as having special expertise in patent law, it is likely that it will bring its existing expertise and framings to bear onto issues of competition.

Indeed, the Federal Circuit is increasingly applying its own substantive law to antitrust issues that implicate patent law, making the problem of coordination particularly pressing. For example, following the Federal

336. See Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 YALE L.J. 1535 (1998); Arti K. Rai, *Specialized Trial Courts: Concentrating Expertise on Fact*, 17 BERKELEY TECH. L.J. 877, 879–80 (2002).

337. See, e.g., Pedraza-Fariña, *supra* note 91.

338. Collins & Evans, *supra* note 135, at 3.

339. See, e.g., Dreyfuss *supra* note 40; Gugliuzza, *supra* note 286, at 1497.

Circuit's decisions in *In re Independent Service Organizations Antitrust Litigation* (Xerox), *Intergraph*,³⁴⁰ and *C.R. Bard*³⁴¹—all of which involved antitrust challenges to a monopolist's refusals to license or sell products subject to intellectual property protection—antitrust attorneys have criticized the Federal Circuit for giving undue weight to intellectual property considerations at the expense of competition principles embedded in antitrust law.³⁴² Even those who defended the Federal Circuit's holdings as consistent with “mainstream antitrust principles,” remarked that the Federal Circuit's antitrust analysis was often “poorly articulated,” “superficial,” “awkward,” and not deeply engaged with the type of “rigorous analysis” required by antitrust law.³⁴³ Importantly, studies of agencies charged with formulating competition policy have found that these agencies tend to downplay intellectual property considerations or fail to consider the impact of breakthrough innovation.³⁴⁴ Taken together, these results are consistent with coordination difficulties predicted from the concentration of expertise in particular communities or institutions. Thus, overcoming coordination difficulties will likely require external incentives or coordination mechanisms.

C. EXTERNAL ADVISORY PANELS AS A PROPOSED SOLUTION TO
TYPECASTING AND COORDINATION PROBLEMS

A key question in the design of expert courts is how to minimize the negative aspects of typecasting and their inability to coordinate—or how to increase the number of frames considered by a community of experts—while maintaining the gains in accuracy derived from expertise. This Part outlines a proposed solution: the creation of advisory panels that would provide a mixture of technical, economic, and sociological expertise by rotating experts in the areas relevant to the case or question under consideration. As a first step, the Federal Circuit could rely on advisory panel opinions when making decisions en banc. The Federal Circuit's most recent en banc orders seek to clarify important areas of the law and involve issues of policy with clear implications for the pace of technological innovation, in which a

340. *Intergraph Corp. v. Intel Corp.*, 253 F.3d 695 (Fed. Cir. 2001); *In re Indep. Serv. Org. Antitrust Litig.*, 203 F.3d 1322 (Fed. Cir. 2000).

341. *C.R. Bard, Inc. v. M3 Sys., Inc.*, 157 F.3d 1340 (Fed. Cir. 1998).

342. See, e.g., Peter Boyle, Penelope M. Lister & J. Clayton Everett, Jr., *Antitrust Law at the Federal Circuit: Red Light or Green Light at the IP-Antitrust Intersection?*, 69 ANTITRUST L.J. 739 (2002).

343. *Id.*

344. See, e.g., Michael L. Katz & Howard A. Shelanski, *Merger Analysis and the Treatment of Uncertainty: Should We Expect Better?*, 74 ANTITRUST L.J. 537, 538 (2007) (noting that antitrust analysis may block mergers with pro-innovation outcomes but allow those with anti-innovation possibilities).

technological, economic, and sociological perspective would be particularly useful.³⁴⁵ Importantly, the Federal Circuit has sought the advice of outside parties in its *en banc* decisions by encouraging the filing of amici briefs.³⁴⁶ The use of advisory panels could gradually be expanded to provide advice in panel decisions that seek to clarify important issues of patent doctrine, and where it would benefit the court to have an opinion on how its decision may impact the pace of technological innovation in a particular field. Advisory panels could also be particularly useful to district courts, when engaging in fact-finding regarding, for example, the level of skill in the art, the predictability of a particular technical field, or the boundaries of analogous arts.

The use of advisory panels with a mixture of economic, sociological, and technological expertise has several advantages over other potential solutions to the CAFC's perceived disengagement from real-world innovation. First, the traditional solutions to the problem of expertise most often used at the trial court level—the use of expert witnesses selected by the parties and the use of court appointed experts—fall short of providing the kind of comprehensive expertise in economic, social, and technical matters that is required to address patent disputes. In addition, parties are expected to find (and often do find) experts who can support their legal conclusions. Litigation turns into a battle of the experts that provides judges and the jury little guidance on how to resolve discrepancies between different experts' perspectives.³⁴⁷ In the context of patent disputes, it is then perhaps no wonder—and not an entirely undesirable outcome—that the Federal Circuit often chooses to substitute its own expertise for that of experts hired by the

345. See, e.g., *Phillips v. AWH Corp.*, 376 F.3d 1382, 1383 (Fed. Cir. 2004) (per curiam) (citations omitted) (order granting petition for rehearing en banc in *Phillips*) (asking for additional amici briefs addressing, among other questions: “What role should prosecution history and expert testimony by one of ordinary skill in the art play in determining the meaning of the disputed claim terms?”); *Ariad Pharmaceuticals, Inc. v. Eli Lilly & Co.*, 595 F.3d 1329, 1330 (Fed. Cir. 2009) (per curiam) (order granting petition for rehearing en banc in *Ariad Pharmaceuticals*, and asking opinions on the proper “scope and purpose of the written description requirement”); Order for rehearing en banc, 2011-1301, *CLS Bank Int'l v. Alice Corp.* (asking: “What test should the court adopt to determine whether a computer-implemented invention is a patent ineligible ‘abstract idea’; and when, if ever, does the presence of a computer in a claim lend patent eligibility to an otherwise patent-ineligible idea?”).

346. *Id.* All of the orders granting a petition for rehearing en banc listed *supra* note 345 specifically invite the PTO or the United States to file amicus briefs and notes that other parties may also file briefs without leave of the court.

347. See Rebecca Haw, *Adversarial Economics in Antitrust Litigation: Losing Academic Consensus in the Battle of Experts*, 106 NW. U.L. REV. 1261, 1270 (2012).

parties. The use of court-appointed, neutral experts may sidestep the battle of the experts, but at the cost of creating a “strong, if not overwhelming, impression of ‘impartiality’ and ‘objectivity,’”³⁴⁸ that may not be warranted. Individual experts likely have pre-existing biases or frames that would not be adequately challenged in the court-appointed expert model. Finally, the Federal Circuit’s increased reliance on amicus briefs, while increasing public participation in important patent law decisions, raises the same concerns of bias as the use of party-appointed experts.

Expert advisory panels avoid these concerns with court-appointed experts, parties-appointed experts, and amici curiae. Expert panels are already widely used to optimize medical decision-making.³⁴⁹ Expert advisory panels can be flexibly designed to reduce pre-existing biases that would be unavoidable with single court-appointed experts. Specifically, studies on the use of expert panels for medical decision-making have shown that reducing bias and increasing accuracy requires balancing a diversity of views and reference frames with panel size.³⁵⁰ Experts can also rotate so as to minimize the possibility of capture, and maximize the fit between expertise and the particular technical, economic, and social problem under study.

Second, advisory panels have several advantages over proposals for decentralization of judicial decision-making. As emphasized above in Section V.A, it will be difficult to assess the effect of different judicial methodologies or framings on innovation, given the national and often international nature of innovative activity—therefore making it hard to ultimately choose among competing judicial approaches. The advisory panels proposal advanced in this Article would ameliorate these concerns because an independent advisory panel would be able to make predictions *ex ante* about the likely economic and social consequences of particular patent policies and rules. In addition, although expanding the number of circuits that can hear patent disputes would certainly ameliorate the dangers of typecasting, it may do so at the expense of expertise. In contrast, advisory panel design can address the typecasting problem (for example, by balancing membership in different specialties and using voting mechanisms that minimize bias) without sacrificing the benefits of expertise. The use of advisory panels, however, does not preclude adopting decentralization proposals. In particular, if an additional circuit court such as the Court of Appeals for the District of

348. *Kian v. Mirro Aluminum Co.*, 88 F.R.D. 351, 356 (E.D. Mich. 1980).

349. See, e.g., Matthew J. Gabel & Charles R. Shipan, *A Social Choice Approach to Expert Consensus Panels*, 23 J. HEALTH ECON. 543, 544 (2004).

350. See, e.g., George A. Krause & James W. Douglas, *Organizational Structure and the Optimal Design of Policymaking Panels*, 57 AM. J. POL. SCI. 135 (2013).

Columbia were to also hear patent cases, as Duffy and Nard propose, advisory panels could provide guidance on patent decisions to both circuits.

Third, as discussed above in Section V.B, proposals to rebalance the Federal Circuit's caseload with a larger proportion of competition cases are likely insufficient to correct coordination problems if the Federal Circuit already views itself as a community with particular expertise in patent law. In addition, coordination problems are not limited to antitrust and patent law—coherent innovation policy should also take into account, for example, the interplay between grants, patents, and other incentives to innovate when considering the impact of patent law on innovation communities.³⁵¹ External advisory panels would be able to more readily take all of these considerations into account and advise the Federal Circuit directly on the impact of particular changes to patent doctrine on the overall innovation landscape.

Fourth, advisory panels can complement Stuart Benjamin and Arti Rai's proposal for the creation of an Office of Innovation hosted within the executive branch that would coordinate the activities of agencies and courts that impact innovation, such as the Federal Communications Commission, the Food and Drug Administration, and the PTO.³⁵² While an office of innovation would be charged with coordinating innovation policy at a macro level, advisory panels could provide input at a micro level. For example, advisory panels can equip the Federal Circuit, and potentially the district courts, with necessary information on the pace of technological innovation, the predictability of a particular field, or the likely impact of market forces on a particular technical advance—information that is often crucial to patentability determinations.

Finally, advisory panels provide the additional advantage of being able to teach and keep the Federal Circuit up to date on developments in innovation economics and sociology, and in the different scientific and technical fields involved in patent litigation. This, in turn, would enhance and standardize the Federal Circuit's expertise—ensuring all of its members are exposed to the most up-to-date insights from research in economics, sociology, and scientific fields relevant to patent law.

An advisory panel opinion on a particular dispute could be made available at the request of any litigant, or at the request of the Federal Circuit itself. To ensure that the Federal Circuit places adequate weight on the panel's recommendations and does not simply ignore its findings, the

351. See, e.g., Daniel J. Hemel & Lisa Larrimore Ouellette, *Beyond the Patents-Prize Debate*, 92 TEX. L. REV. 303 (2013).

352. Benjamin & Rai, *supra* note 281, at 56–58.

recommendations should be made part of the record, and the Federal Circuit should be bound to clearly explain its reasoning if it chooses to deviate from them. This hard-look requirement would ensure that the Federal Circuit retains flexibility in adopting the panel's opinions but is also bound to consider them seriously.³⁵³ Section 175 of the Patent Act already provides for the appointment of technical assistants to the Federal Circuit.³⁵⁴ The drafters of the Act characterized the role of such assistants as “confer[ring] with judges on technical as well as legal matters.”³⁵⁵

This proposal assumes that there are real gains from expert decision-making in patent law—conceptualized as substantial expertise—that can be preserved while minimizing the costs derived from typecasting and the inability to coordinate. Both defending that assumption and providing a full answer to how to optimize Federal Circuit expertise, including a full blueprint for designing advisory panels, requires further research. In particular, it requires an in-depth exploration of the kinds of substantive expertise involved in the resolution of patent disputes, and how to best achieve the correct mix of expertise in panel design.³⁵⁶ For example, there are at least five types of possible expertise relevant to patent law: (1) expertise in the science and technology involved in the discovery, (2) expertise in innovation dynamics, both sociological and economic, (3) expertise in patent law through continued exposure to relevant cases, or through previous practice experience, (4) expertise in complex litigation, and (5) meta-expertise at the intersection of these four types of expertise. Because addressing patent law problems requires reliance on knowledge at the intersection of different fields, bringing together the right types of expertise to solve patent disputes will also face coordination challenges similar to those in assembling medical review panels or policy panels. Operationalizing advisory panels will also require developing protocols for member selection, creating incentives for both the Federal Circuit and parties to actually use expert panels,³⁵⁷ and

353. See Benjamin & Rai, *supra* note 281, at 65 (proposing the creation of an executive “Office of Innovation Policy” whose opinions—while not legally binding—would “qualify as material at which the agency should take a hard look, and to which the agency would be required to respond”).

354. 28 U.S.C. § 175 (2012).

355. S. Rep. No. 97-275, at 17 (1981).

356. See, e.g., Gabel & Shipan, *supra* note 349.

357. See, e.g., Christopher Tarver Robertson, *Blind Expertise*, 85 N.Y.U. L. REV. 174, 200–01 (2010) (noting judges rarely use court-appointed experts because of social norms against their use and the high cost of finding and training experts to testify). To address hurdles in the assembly of expert panels, Rule 706 and Section 175 could be amended to encourage specialized or expert courts to appoint expert advisory panels, or to encourage the

evaluating whether expert panel opinions should also be available to advise district courts and agencies such as the PTO.³⁵⁸

Given the main focus of this Article on modeling Federal Circuit behavior as a function of expert community dynamics, its treatment of advisory panels design is necessarily in outline form. A full analysis of how expert panels would be designed and implemented will require an in depth exploration of the types of substantive expertise needed to make decisions in patent law, an analysis of how advisory panels have been put together in other disciplines and whether those insights can be adapted to patent law, as well as elucidating details regarding membership selection criteria, mechanisms by which the parties could challenge expert reports, and the scope of the panel's mandate. In addition, a key question regarding advisory panel design is whether—rather than provide technical advice to the Federal Circuit—they are best housed within trial courts (where they would function to provide expert guidance on factual issues), or the PTO. For example, given the PTO's expanding role in patent law following the passage of the America Invents Act, advisory panels could also (or alternatively) be available to provide expert advice on appeals to the Patent Trial and Appeal Board.³⁵⁹

VI. CONCLUSION

The Federal Circuit sits at the epicenter of a vigorous debate over the role of specialized courts in a broader system of generalist judges. Critics of the Federal Circuit view the court as a failed experiment in judicial specialization. They point to its over-reliance on inflexible rules, its refusal to accord deference to both district courts and the PTO, and its failure to maintain doctrinal uniformity. In contrast, supporters of judicial specialization in patent law warn that decentralization would lead to increased forum-shopping and a high level of uncertainty regarding the applicable legal regime—ultimately dampening innovation.

This Article has argued that any principled discussion of Federal Circuit design requires an understanding of decision-making by expert communities.

Federal Circuit to do so. Also, professional organizations of scientists, economists, and sociologists could be enlisted to provide help in panel assembly.

358. A growing literature analyzes whether and how institutions other than the Federal Circuit should be involved in the development of patent law. *See, e.g.,* Arti Rai, *Patent Validity Across the Executive Branch: Ex Ante Foundations for Policy Development*, 61 DUKE L.J. 1237 (2012); Dreyfuss, *supra* note 10. Advisory panels could thus be useful to a variety of institutional players.

359. *See* Dreyfuss, *supra* note 10.

In particular, it requires addressing three fundamental questions: First, how does subject-matter specialization or expertise impact the content of judicial decisions? Second, how does subject-matter specialization or expertise impact the form of judicial decisions? Finally, how does subject-matter specialization or expertise impact the relationship between decision-making bodies? Drawing on a rich literature on the sociology of expertise, this article takes a first step in answering these key questions by developing a typology of four features of decision-making by expert communities. The Article demonstrates how these features explain puzzling aspects of Federal Circuit jurisprudence, such as lack of deference to both the district courts and the PTO, and a preference for rules over standards, and predicts additional behaviors such as defiance of the Supreme Court, jurisdictional expansion, and inconsistent rule application.

Importantly, the typology has two broader implications. First, it identifies two specific features of expert communities—typecasting and an inability to self-coordinate—that are normatively undesirable in the context of a centralized expert court. This Article analyzes existing proposals for Federal Circuit reform, concluding that most of them are unable to address distortions in decision-making caused by these two features of an expert community's behavior. As an alternative, this Article proposes the adoption of expert advisory panels with the necessary complement of economic, social, and technological expertise. Second, it develops a novel theoretical lens with implications for analyzing the behavior of other expert courts. Further studies analyzing the behavior of, for example, bankruptcy and tax courts can both test and refine the model for application to these broader contexts.

SOFTWARE COPYRIGHT'S ORACLE FROM THE CLOUD

Lothar Determann[†] & David Nimmer^{††}

ABSTRACT

Clouds are on the horizon for software copyrights. The open source movement is actively trying to turn copyright into “copyleft.” Courts around the world are reshaping the first sale doctrine, notably the European Court of Justice in *UsedSoft v. Oracle* not to mention the United States Court of Appeals for the Ninth Circuit in *Autodesk v. Vernor*. Software manufacturers are fleeing from distribution models toward service models—into the cloud. A perfect storm for software copyrights is brewing. The cloud promises to enable software publishers to place their code outside both the framework of copyright exhaustion under the first sale doctrine and the “distribution trigger” in open source code license terms. Users’ inability, in the cloud context, to directly access the underlying software threatens to exert various side effects, notably affecting software interoperability. New kids on the block lose the ability to reverse engineer hosted software. Established platform providers gain the ability to prevent interoperability, based on laws prohibiting interference with computers and technical protection measures. These developments risk upsetting the delicate balance between exclusive rights for copyright owners and access privileges for the public—a balance that courts and legislatures have carefully established over the years—in order to foster creativity and innovation. With unprecedented pressure on traditional distribution models, how will copyright law cope? In our Article, we attempt to illuminate the immediate path ahead, discuss possible answers, and pose more questions.

© 2015 Lothar Determann and David Nimmer. Opinions expressed herein reflect only the authors’ views, and should not be imputed to their universities, firms, clients, or others.

[†] Lothar Determann teaches computer, internet, and data privacy law at Freie Universität Berlin; University of California, Berkeley, School of Law; and Hastings College of the Law, San Francisco and practices technology law as a partner with Baker & McKenzie LLP, admitted in California and Germany.

^{††} David Nimmer teaches domestic and international copyright at the UCLA School of Law and as a guest at various other institutions. He has represented a wide array of clients in copyright matters at Irell & Manella LLP.

TABLE OF CONTENTS

I.	INTRODUCTION	163
II.	HISTORY AND APPLICATION OF COPYRIGHT PROTECTION FOR SOFTWARE	165
A.	CONFERRAL OF RIGHTS ON COPYRIGHT PROPRIETORS	165
1.	<i>Titration of the appropriate amount of copyright protection</i>	165
2.	<i>According protection to software's artistry, not its functionality</i>	166
3.	<i>Computer Software Amendment of 1980</i>	167
4.	<i>International dimension</i>	171
B.	RESERVATION OF PRIVILEGES TO SOFTWARE USERS.....	172
1.	<i>First sale doctrine</i>	172
2.	<i>Essential step defense</i>	173
3.	<i>Reverse engineering</i>	175
C.	DIVERGENT INTERPRETATIONS OF SOFTWARE COPYRIGHT IN AND BETWEEN THE UNITED STATES AND THE EUROPEAN UNION	176
1.	<i>Starting in the United States</i>	176
2.	<i>Moving to Europe</i>	180
3.	<i>Cross-breezes: exceptions in United States and E.U. to the paradigms set forth above</i>	183
a)	Returning Stateside	184
b)	Final European Jaunt.....	188
4.	<i>Departure from the statute's goals?</i>	189
III.	ECONOMICS AND TECHNOLOGY OF THE CLOUD	193
A.	SOFTWARE COMMERCIALIZATION MODELS	193
B.	EXAMPLES OF CLOUD OFFERINGS	195
C.	MULTIPLE CLOUD USERS.....	196
IV.	THE CLOUD'S IMPACT ON SOFTWARE COPYRIGHTS	201
A.	CHALLENGES TO COPYRIGHT LAW'S GOALS	201
1.	<i>First sale doctrine</i>	201
2.	<i>Essential step defense</i>	203
3.	<i>Reverse engineering</i>	203
B.	OTHER IMPACTS	205
1.	<i>Open-source-license terms</i>	206
2.	<i>Unauthorized Access to Copyrighted Material</i>	206
a)	Piracy.....	206
b)	Facilitation of Multiple Cloud Users	207
V.	SOME WEATHER FORECASTS	208

A.	GLOOMY PROSPECTS: DARK CLOUDS UNDERMINE THE GOALS OF 1980	208
B.	A CHEERY REBUTTAL: RETURN TO AN EARLIER SENSIBILITY	210
C.	HIGHS AND LOWS IN TURBULENT TRANSITIONS	211

I. INTRODUCTION

It was never a happy marriage. Computer software and copyright law came from different families, embodied disparate values, pursued separate interests. So, even though the latter became the vehicle to secure legal protection for the former, the union was never smooth—even on the wedding day. At present, the plates are flying with ever more reckless abandon. As the open source movement has been active to turn copyright into “copyleft,”¹ courts around the world are reshaping the first sale doctrine, and software manufacturers are fleeing from distribution to service models, and most recently, into the cloud. With a perfect storm brewing, this Article steps in to offer some mild marital counseling.

Part of the initial bargain that brought software within the copyright fold was that users would obtain privileges in software via two legal devices—the “first sale” defense and the “essential step” doctrine. Yet, in recent years, software developers have striven to maximize their rights by enclosing their physical products with shrinkwrap or other purported contractual terms. Through that stratagem, they have sought to avoid the privileges that would otherwise inure to software users via the first sale doctrine. In that endeavor, they have largely achieved success within the United States, albeit not without some backlash.² Yet, in *UsedSoft v. Oracle*, the European Court of Justice dealt Oracle and other software makers a blow in July 2012 by dramatically expanding the scope of “exhaustion,” the European analog to the United States’ first sale doctrine.³ Thus, the international status of these efforts is currently in flux.⁴ Moreover, even within those two jurisdictions, the results have been less than categorical—for instance, in March 2013, the United States Supreme Court held that copies made and first sold abroad can be imported into the United States against the U.S. copyright owner’s will,

1. Lothar Determann, *Dangerous Liaisons—Software Combinations as Derivative Works? Distribution, Installation, and Execution of Linked Programs Under Copyright Law, Commercial Licenses, and the GPL*, 21 BERKELEY TECH. L.J. 1421, 1481 (2006).

2. See discussion *infra* Subsection II.C.1.

3. Case C-128/11, *UsedSoft GmbH v. Oracle Int’l Corp.*, 2012 EUR-Lex CELEX, ¶¶ 44–46 (July 3, 2012) (discussed *infra* Subsection II.C.2); see also GREGORY J. BATTERSBY ET AL., 2014 LICENSING UPDATE §3.01, at 3-3 (2014).

4. See discussion *infra* Section IV.A

potentially presaging a retreat from the expansive interpretation of “license” terms evident in prior opinions.⁵ Correlatively, more recent developments within Europe call into question the European Court of Justice ruling of 2012—some portending its limitation, others its radical expansion.⁶

As if those developments were not dizzying enough in their own right, the advent of the cloud with service-based—as opposed to distribution-based—commercialization models⁷ injects even more volatility into the mix. By keeping software copies on their own servers and making software functionality available remotely as part of cloud offerings on a subscription basis, the cloud enables software publishers to place their code outside the framework of copyright exhaustion under both the first sale doctrine and the “distribution trigger” in open source code license terms.⁸ Users’ inability, in the cloud context, to directly access the underlying software threatens to exert various side effects, notably affecting software interoperability, and new kids on the block lose the ability to reverse engineer hosted software.⁹ Established platform providers gain the ability to prevent interoperability, based on laws prohibiting interference with computers and technical protection measures.

These developments risk upsetting the delicate balance between exclusive rights for copyright owners and access rights for the public, a balance that courts and legislatures have carefully established over the years in order to foster creativity and innovation. With unprecedented pressure on traditional distribution models, how will copyright law cope? Having no idea of the ultimate goal towards which the law is heading, we attempt to illuminate the immediate path ahead, in the light of past events that have brought matters to their current stance.

5. See *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351 (2013) (discussed *infra* Subsection II.C.3.a)).

6. See discussion *infra* Part II.C.3.b)

7. See *infra* Part III.

8. Most open source license terms allow running and using software without significant restrictions and apply license conditions only if and when the licensee starts distributing the code or modifications thereof; see, for example, Section O, paragraph 2 of the GPLv2, FREE SOFTWARE FOUNDATION (June 1991), <http://www.gnu.org/licenses/gpl-2.0.html>; Lothar Determann, Stuart Pixley & Gary Shapiro, *Managing commercial risks in open source software licensing*, J. INTELL. PROP. L & PRAC., 2007, p. 1, 6.

9. Regarding interoperability, see discussion *infra* Subsection IV.A.3; regarding reverse engineering, see discussion *infra* Subsection II.B.3

II. HISTORY AND APPLICATION OF COPYRIGHT PROTECTION FOR SOFTWARE

From the beginnings of computer technology, copyright law was an unlikely candidate to protect intellectual property in software. After all, copyright is intended, at base, to protect creative expression, such as paintings, music, novels, and sculptures. Yet the value in most software inheres not in its creative expression, but instead in its technical functionality—a matter that copyright law was never intended to protect. In fact, the United States Copyright Act expressly carves out ideas, procedures, processes, systems and methods of operation from its scope of protection.¹⁰

Nevertheless, copyright law ended up as the legal vehicle providing the primary source of intellectual property law protection for software.¹¹ It took considerable bending and tweaking to make it work.¹² Even the rise of open source licensing and “copyleft”¹³ can be viewed as an affirmation of software copyright law, insofar as copyleft and the broader trend of “sharing nicely”¹⁴ were enabled precisely through leveraging the adaptation right afforded to copyright owners—a feature that patent and trade secret laws pointedly do not offer.¹⁵

A. CONFERRAL OF RIGHTS ON COPYRIGHT PROPRIETORS

1. *Titrating the appropriate amount of copyright protection*

As contemplated by the U.S. Constitution,¹⁶ the Copyright Act protects creative works through exclusive rights that encourage further creation and public availability of such works.¹⁷ Those exclusive rights enable owners to permit or prohibit reproductions and other exploitations of their works,¹⁸

10. 17 U.S.C. § 102(b) (2012).

11. See Peter S. Menell, *Tailoring Legal Protection for Computer Software*, 39 STAN. L. REV. 1329, 1354 (1987).

12. See *id.* at 1329; see also Mark A. Lemley, *Convergence in the Law of Software Copyright?*, 10 BERKELEY TECH. L.J. 1, 3–6 (1995); Pamela Samuelson et al., *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308 (1994).

13. Determann, *supra* note 1, at 1481.

14. Yochai Benkler, *Sharing Nicely: On Shareable Goods and the Emergence of Sharing as a Modality of Economic Production*, 114 YALE L.J. 273, 334–36 (2004).

15. Determann, *supra* note 1, at 1442.

16. U.S. CONST. art. I, § 8, cl. 8 provides that Congress shall have the power “[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”

17. See Pamela Samuelson, *Computer Programs, User Interfaces, and Section 102(b) of the Copyright Act of 1976: A Critique of Lotus v. Paperback*, 55 LAW & CONTEMP. PROBS. 311, 339 (1992).

18. The full roster of exclusive rights is contained in 17 U.S.C. § 106 (2012).

charging permission fees to utilize the work. The prospect of such license charges are intended to incentivize authors to create and to adapt original works.

But the public interest in creative works may be derogated by either under- or over-protection.¹⁹ If facts and functionality could be copyrighted by the first author who writes about them, for example, then later authors could not create their own creative works of authorship based on the same facts or functionality. If the first author who writes a book about an accounting method refuses to grant licenses, others could not write better books about the same method.²⁰ On the other hand, if an author could not achieve some measure of protection against copying of her book about accounting methods, then perhaps no one would wish to expend the labor in the first place to write one. Accordingly, legislatures and courts have, over the years, endeavored to maintain a delicate balance²¹ between granting exclusive rights to authors and affording access privileges to the public.²²

2. *According protection to software's artistry, not its functionality*

As the vector on which that balancing act is performed, the Copyright Act attentively limits copyrightable subject matter to creative, artistic expression by preserving underlying ideas and functionality in the public domain:

In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work.²³

Historically speaking, this express exception made software particularly ill-suited for copyright protection—one need only consider that the value of

19. See Samuelson, *supra* note 17, at 338–40; see also Determann, *supra* note 1, at 1438.

20. See *Baker v. Selden*, 101 U.S. 99, 107 (1879).

21. A paean to that delicate balance fills David Nimmer, Elliot Brown & Gary N. Frischling, *The Metamorphosis of Contract into Expand*, 87 CALIF. L. REV. 17 (1999). Many of the themes from that Article carry forward here.

22. See, e.g., *Eldred v. Ashcroft*, 537 U.S. 186 (2003) (holding extended copyright terms constitutional); *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984) (ruling that marking individual copies of complete television shows for the purposes of time-shifting constitutes permissible fair use, not violation of copyright); *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539 (1985) (holding that fair use is not an appropriate defense for using excerpts from the unpublished memoirs of President Gerald Ford). Where this line should be drawn has been the subject of extensive academic debate, especially in the last decade. See, e.g., Determann, *supra* note 1.

23. 17 U.S.C. § 102(b) (2012).

most computer programs²⁴ lies in their functionality and efficiency.²⁵ Enterprise and office software users typically appreciate, and pay for, the speed, reliability, and operational simplicity with which a particular program produces results. The simpler, more common, and more familiar a graphic user interface (“GUI”), the better.²⁶ Most software users fail even to notice, and certainly care little or nothing at all for, creative flourishes embedded in the software.²⁷

3. *Computer Software Amendment of 1980*

Copyright’s dominion over software has long been settled in the United States. Through its amendment of Title 17 of the United States Code (containing the pre-existing Copyright Act) to include the Computer Software Copyright Act of 1980,²⁸ Congress²⁹ chose copyright law as the

24. Video and computer games and other entertainment software are noteworthy exceptions.

25. David G. Luetggen, *Functional Usefulness vs. Communicative Usefulness: Thin Copyright Protection for the Nonliteral Elements of Computer Programs*, 4 TEX. INTELL. PROP. L.J. 233, 249–60 (1996); *see also* Determann, *supra* note 1, at 1438–39.

26. *See* Lotus Dev. Corp. v. Borland Int’l, Inc., 49 F.3d 807 (1st Cir. 1995), *aff’d by an equally divided Court*, 516 U.S. 233 (1996) (holding a menu hierarchy not copyrightable and citing concerns with interoperability and unduly burdening customers to learn multiple systems).

27. For a co-author’s in-depth discussion of this dissonance, *see* Determann, *supra* note 1, at 1438–42.

28. Act of Dec. 12, 1980, Pub. L. No. 96-517, § 10, 94 Stat. 3028.

29. Congress acted against the backdrop of a special commission, the National Commission on New Technological Uses of Copyrighted Works (“CONTU”), that it had charged with investigating this arena. In his CONTU concurrence, Commissioner and Vice-Chair Melville Nimmer issued a warning that will be reprised later in this Article:

At the same time I should like to suggest a possible line of demarcation, which would distinguish between protect[a]ble and nonprotect[a]ble software in a manner more consistent with limiting such protection to the conventional copyright arena. This suggestion is made not because I recommend its immediate implementation, but rather because it may prove useful in the years to come if the Commission’s recommendation for protection of all software should prove unduly restrictive. In such circumstances it may prove desirable to limit copyright protection for software to those computer programs which produce works which themselves qualify for copyright protection. A program designed for use with a data base, for example, would clearly be copyrightable since the resulting selection and arrangement of items from such data base would itself be copyrightable as a compilation.

CONTU, FINAL REPORT OF THE NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS 27 (1979); *see* David Nimmer, *Codifying Copyright Comprehensibly*, 51 UCLA L. REV. 1233, 1263–65 (2004).

vehicle for courts to afford protection to computer programming.³⁰ Computer programs (both object and source code versions) therefore now generally find protection under the U.S. Copyright Act.³¹ The underlying code constitutes a literary work.³² Output of software, such as screen displays and GUIs, do not constitute software in and of themselves, but can be protected separately from underlying code as literary, pictorial or graphic works.³³

Copyright law is designed to protect creative expression as an incentive for further creative activity. In line with that goal, the U.S. Supreme Court has unanimously rejected copyright protection for mere “sweat of the brow.”³⁴ A paradox arises in that software’s value is usually measured precisely by its functionality and efficiency,³⁵ aspects expressly excluded from copyright protection.³⁶ In light of that disconnect, judges are faced with the

30. See, e.g., *Apple Computer, Inc. v. Franklin Computer Corp.*, 714 F.2d 1240 (3d Cir. 1983) (finding that a computer program is a literary work and protected under copyright); for an overview, see MARK A. LEMLEY, PETER S. MENELL, ROBERT P. MERGES & PAMELA SAMUELSON, *SOFTWARE AND INTERNET LAW* 1–45, 97–98 (1st ed. 2000), 33–35 (2d ed. 2003).

31. See, e.g., *Williams Elecs., Inc. v. Artic Int’l, Inc.*, 685 F.2d 870 (3d Cir. 1982); *Franklin Computer*, 714 F.2d at 1240.

32. *Franklin Computer*, 714 F.2d at 1246–49.

33. See *Stern Elecs., Inc. v. Kaufman*, 669 F.2d 852, 855–57 (2d Cir. 1982) (holding that video game manufacturers could copyright image and sounds in a game, not just the underlying source code); *Data E. USA, Inc. v. Epyx, Inc.*, 862 F.2d 204 (9th Cir. 1988); *Apple Computer, Inc. v. Microsoft Corp.*, 35 F.3d 1435 (9th Cir. 1994) (addressing the scope of protection for GUI elements).

34. See *Feist Publ’n, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 352–60 (1991).

35. See *Bucklew v. Hawkins, Ash, Baptie & Co.*, 329 F.3d 923, 928 (7th Cir. 2003) (finding software whose primary function was improving efficiency copyrightable).

36. *Computer Assocs. Int’l, Inc. v. Altai, Inc.*, 982 F.2d 693, 712 (2d Cir. 1992); see also Mark A. Lemley & David W. O’Brien, *Encouraging Software Reuse*, 49 STAN. L. REV. 255 (1997) (arguing that copyright protection for software inhibits efficiency); Peter S. Menell, *An Analysis of the Scope of Copyright Protection for Application Programs*, 41 STAN. L. REV. 1045 (1989) (arguing that viewing application programming as literary creativity rather than a technological advancement may give too broad a protection); Samuelson et al., *supra* note 12 (arguing that law should focus on value and know-how rather than means which might be used to appropriate it). Cf. Jane C. Ginsburg, *Four Reasons and a Paradox: The Manifest Superiority of Copyright Over Sui Generis Protection of Computer Software*, 94 COLUM. L. REV. 2559 (1994) (noting that while imperfect, copyright is preferable to other regimes for protecting software); Arthur R. Miller, *Copyright Protection for Computer Programs, Databases, and Computer-Generated Works: Is Anything New Since CONTU?*, 106 HARV. L. REV. 977 (1993) (defending CONTU as a flexible regime to address current and future issues in software copyright); Stacey L. Dogan & Joseph P. Liu, *Copyright Law and Subject Matter Specificity: The Case of Computer Software*, 61 N.Y.U. ANN. SURV. AM. L. 203, 204 (2005) (discussing judicial treatment of computer software copyright protection in light of its dual nature).

prospect of having to fit a square peg into a round hole.³⁷ Courts have developed a number of tests and approaches to separate software's protectable creative elements from its non-protectable functional elements.³⁸ Creative elements are protected against literal and non-literal copying. By contrast, functional ingredients can be freely duplicated, particularly when idea and expression (or functionality and creativity) merge or when expression is dictated by external requirements (*scènes à faire*—for example, a multi-tier technical solution in which certain individual tiers can only be programmed efficiently in one particular manner).³⁹ Thus, software copyright analysis must begin with a thorough examination of the dichotomy between creativity and functionality.⁴⁰ This analysis involves filtering out ideas, processes, methods, facts, and elements dictated by external factors or efficiency; material in the public domain; expression which has merged with any of the foregoing; and expression which is so standard or common as to be a “necessary incident” to any of the foregoing.⁴¹

Software can present questions at multiple levels. Artistic screen displays of computer games (with fantasy figures and landscapes)⁴² bear a greater resemblance to traditional subjects of copyright protection (a novel, a still-life) than to software in executable form⁴³ or to the functionality-driven user

37. See *Altai*, 982 F.2d at 712 (“Generally, we think that copyright registration—with its indiscriminating availability—is not ideally suited to deal with the highly dynamic technology of computer-science.”).

38. *Id.* at 714; *Gates Rubber Co. v. Bando Chem. Indus., Ltd.*, 9 F.3d 823, 836–38 (10th Cir. 1993); *Lotus Dev. Corp. v. Borland Int'l*, 49 F.3d 807, 815–17 (1st Cir. 1995), *aff'd by an equally divided Court*, 516 U.S. 233 (1996); *Chamberlain Grp., Inc. v. Skylink Techs., Inc.*, 381 F.3d 1178, 1191–92 (Fed. Cir. 2004); *Lexmark Int'l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522, 534–36 (6th Cir. 2004); *Case C-406/10, SAS Institute Inc. v. World Programming Ltd.*, 2012 E.C.R. I-0000, at ¶ 46.

39. See *Oracle Am., Inc. v. Google Inc.*, 872 F. Supp. 2d 974 (N.D. Cal. 2012), *rev'd*, 750 F.3d 1339 (Fed. Cir. 2014); *Altai*, 982 F.2d at 707–08; *Case C-406/10, SAS Inst. Inc. v. World Programming Ltd.*, 2012 E.C.R. I-0000, at ¶ 46.

40. See *Feist Publ'n's, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 362 (1991); *Bateman v. Mnemonics, Inc.*, 79 F.3d 1532, 1543–45 (11th Cir. 1996); *Apple Computer, Inc. v. Microsoft Corp.*, 35 F.3d 1435, 1443–46 (9th Cir. 1994); *Eng'g Dynamics, Inc. v. Structural Software, Inc.*, 26 F.3d 1335, 1342–47 (5th Cir. 1994); *Kepner-Tregoe, Inc. v. Leadership Software, Inc.*, 12 F.3d 527, 534 (5th Cir. 1994); *Gates Rubber*, 9 F.3d at 834–37; *Atari Games Corp. v. Nintendo of Am., Inc.*, 975 F.2d 832, 839 (Fed. Cir. 1992); *Control Data Sys., Inc. v. Infoware, Inc.*, 903 F. Supp. 1316, 1322 (D. Minn. 1995); *CMAX/Cleveland, Inc. v. UCR, Inc.*, 804 F. Supp. 337, 352–53 (M.D. Ga. 1992); *Brown Bag Software v. Symantec Corp.*, 960 F.2d 1465, 1475–76 (9th Cir. 1992); *MiTek Holdings, Inc. v. Arce Eng'g Co.*, 89 F.3d 1548, 1554–56 (11th Cir. 1996).

41. *Gates Rubber*, 9 F.3d at 833, 837–38.

42. See, e.g., *Micro Star v. Formgen Inc.*, 154 F.3d 1107, 1111–12 (9th Cir. 1998).

43. Machine code is not human-readable, consisting of simply “on” or “off” states, typically represented by zeros and ones.

interfaces found in application programs.⁴⁴ Courts have defined the dividing line on a case-by-case basis in light of the underlying public policy considerations, often disagreeing on methods and results. An example is the epic battle that recently unfolded between Oracle and Google over Google's use of Java application programming interface ("API") packages.⁴⁵ Whereas the district court had denied copyright protection to the headers, sequence, and organization of the API packages on a number of grounds,⁴⁶ the Federal Circuit Court of Appeals reversed.⁴⁷ Assuming that appellate court determination survives, future defendants will have to prove fair use, merger, or *scènes à faire* in order to vindicate copying of interfaces, lock-out codes, and other gateways to interoperability.⁴⁸

Some cases have tested copyright law's boundaries with respect to highly valuable aspects of software that were not themselves creative, and hence outside the bounds of copyright protection.⁴⁹ These situations involved companies using otherwise creative works in a purely functional manner, e.g., as interfaces, passwords, or lock-out mechanisms. Courts denied copyright protection for such works, regardless of how creative and original they were, in the interest of preserving the balance between protection and access rights described above.⁵⁰ Thus, when proprietors employ copyrighted works in a

44. *See, e.g.*, *Lotus Dev. Corp. v. Borland Int'l, Inc.*, 49 F.3d 807 (1st Cir. 1995), *aff'd by an equally divided Court*, 516 U.S. 233 (1996); *Computer Assocs. Int'l, Inc. v. Altai, Inc.*, 982 F.2d 693, 715 (2d Cir. 1992).

45. *Oracle*, 872 F. Supp. 2d 974, *rev'd*, 750 F.3d 1339 (Fed. Cir. 2014).

46. 872 F. Supp. 2d at 998.

47. *Oracle*, 750 F.3d 1339, 1361 (Fed. Cir. 2014).

48. *Id.* at 1358. The court indicated that it would be difficult for Google to substantiate any of these defenses, based on its determination that Google did not create Android as interoperable with Java, and that Google adopted the API declaring code, organizational sequence, and structures as a mere convenience for developers who were used to the Java environment. *Id.* at 1371. But, regarding fair use, the court ultimately remanded back to the district court. *Id.* at 1373–77. In the interim, the Supreme Court expressed interest in reviewing the Federal Circuit's ruling by inviting the Solicitor General to express the views of the United States in the case. 135 S. Ct. 1021 (2015). The latter urged against granting review, in the process rejecting Google's view that the APIs in suit amounted to a "method of operation." Brief for the United States as Amicus Curiae at 14, No. 14-410 (May 26, 2015).

49. *See, e.g.*, *Midway Mfg. v. Artic Int'l, Inc.*, 704 F.2d 1009, 1014 (7th Cir. 1983) ("[T]he amount by which the language of Section 101 must be stretched to accommodate speeded-up video games is, we believe, within the limits within which Congress wanted the new Act to operate.").

50. *Oracle Am., Inc. v. Google Inc.*, 810 F. Supp. 2d 1002 (N.D. Cal. 2011), *rev'd*, 750 F.3d 1339 (Fed. Cir. 2014) (on technical user interfaces); *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1521 (9th Cir. 1992) (holding that reverse engineering to obtain interface information for compatibility is permissible under narrow circumstances); *Chamberlain Grp., Inc. v. Skylink Techs., Inc.*, 381 F.3d 1178 (Fed. Cir. 2004); *Lexmark Int'l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522 (6th Cir. 2004).

software context for purposes apart from its creative expression, they cannot uncritically rely on the protections copyright law affords.⁵¹

4. *International dimension*

Copyright law has established itself globally as the primary intellectual property regime for software. Jurisdictions that initially rejected this approach have since adopted it.⁵² In the thirty-one member states of the European Economic Area (“EEA”),⁵³ for example, national copyright law is partially harmonized by Directives of the European Union, in particular the E.U. Software Directive⁵⁴ and the E.U. Copyright Directive.⁵⁵ The E.U. Software Directive grants copyright protection to computer programs as literary works.⁵⁶ Like U.S. copyright law,⁵⁷ the E.U. Software Directive protects only creative elements of computer programs and not functionality, technical interfaces, programming language or data file formats.⁵⁸ As such, the E.U.

51. See *Sega*, 977 F.2d 1510; *Lexmark*, 387 F.3d 522. In *Oracle America, Inc. v. Google Inc.*, the Federal Circuit took a narrow view of interoperability and did not accept the adoption of familiar headers or organizational structures for convenience purposes. 750 F.3d 1339, 1361 (Fed. Cir. 2014).

52. An example is Germany: Software had been expressly recognized in Section 2 of the German Copyright Act as a category of copyrightable works since 1985. Urheberrechtsgesetz, UrhG [Copyright Law], Sept. 9, 1965, as amended May 8, 1998 § 2 no. 2 (F.R.G.) (Ger.). However, prior to the implementation of the EC Software Directive into German law in 1993, German courts had required a very high level of originality before they would afford copyright protection for software. The leading cases are from 1985 and 1991: Bundesgerichtshof [BGH] [Federal Court of Justice] May 19, 1985, 12 Gewerblicher Rechtsschutz und Urheberrecht [GRUR] 1041 (F.R.G.) (known as the collection program case—“Inkasso-Programm”) and Bundesgerichtshof [BGH] [Federal Court of Justice] 1991, 19 Neue Juristische Wochenschrift [NJW] 1231 (F.R.G.) (Ger.) (known as the operating system case—“Betriebssystem”). Many programs that would have easily qualified as copyrightable in the United States were not found to be so in Germany. For an overview of software copyright protection in the European union, see Pamela Samuelson, *Comparing U.S. and EC Copyright Protection for Computer Programs: Are They More Different Than They Seem?*, 13 J.L. & COM. 279 (1994).

53. The EEA consists of the 28 EU Member States, which includes Croatia since July 1, 2013, plus Norway, Iceland, and Liechtenstein.

54. Directive 2009/24/EC, of the European Parliament and of the Council of April 23, 2009 on the Legal Protection of Computer Programs, 2009 O.J. (L 111) 16.

55. Directive 2001/29/EC, of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society, 2001 O.J. (L 167) 10.

56. Directive 2009/24/EC, *supra* note 54, art. 1(1).

57. 17 U.S.C. §§ 102(a)–(b) (2012).

58. Directive 2009/24/EC, *supra* note 54, art. 1(2); Case C-406/10, SAS Inst. Inc. v. World Programming Ltd., 2012 E.C.R. I-0000, at ¶ 46.

Software Directive does not cover GUIs.⁵⁹ Ordinary copyright law can cover a GUI, but only if and to the extent the GUI is sufficiently original and not merely dictated by functional requirements.⁶⁰ Unlike the code that creates the GUI, the GUI itself does not constitute a computer program.⁶¹

B. RESERVATION OF PRIVILEGES TO SOFTWARE USERS

1. *First sale doctrine*

Congress's decision to embrace software within the framework of Title 17 by the Computer Software Copyright Act of 1980 inevitably brought innumerable features of copyright doctrine to the software legal environment. Had Congress chosen a different vehicle, there would have been no importation into the software domain of such incidents of pre-existing copyright law as the necessity to place a © on computer code;⁶² it would have been unnecessary to register software for protection with the United States Copyright Office; and its infringement need not have been gauged through a standard called "substantial similarity,"⁶³ in turn calibrated by another doctrine called "fair use."⁶⁴

Among the many other antecedent doctrines of copyright law imported into the software context by virtue of the 1980 amendment is the "first sale" doctrine.⁶⁵ More than a century ago, in *Bobbs-Merrill Co. v. Straus*, the U.S. Supreme Court ruled that a book publisher could not use copyright law to enforce minimum resale price covenants against secondary book purchasers, inasmuch as the copyright owner exhausted his right to control distribution after the first sale of each book.⁶⁶ That ruling about the very forgettable novel *The Castaway* gave birth to a vital feature of copyright law, the first sale doctrine, which was subsequently adopted into legislation and treaties around

59. Case C-393/09, *Bezpečnostní softwarová asociace-Svaz softwarové ochrany v. Ministerstvo kultury*, 2010 E.C.R. I-13971, at ¶ 42.

60. *Id.* at ¶¶ 49, 51.

61. *Id.*

62. Of course, in 1989, incident to joining the Berne Convention, the severity of that requirement was greatly alleviated. See 2 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 7.02 (2013).

63. See *Computer Assocs. Int'l, Inc. v. Altai, Inc.*, 982 F.2d 693, 715 (2d Cir. 1992). See generally David Nimmer, Richard L. Bernacchi & Gary N. Frischling, *A Structured Approach to Analyzing the Substantial Similarity of Computer Software in Copyright Infringement Cases*, 20 ARIZ. ST. L.J. 625 (1988).

64. See 4 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.03[F][5] (2013); Nimmer, Bernacchi & Frischling, *supra* note 63, at 651–56.

65. For a more detailed background on doctrine's origin, see *Used Software Sales and Copyright Exhaustion in Europe*, 17 BNA ELEC. COM. & L. REP. 2149 (2012).

66. See *Bobbs-Merrill Co. v. Straus*, 210 U.S. 339 (1908).

the world as part of the overall effort to balance exclusion and access rights. Many aspects of the first sale doctrine appear relatively uncontroversial today: for books and many other products, the copyright owner exhausts any exclusive distribution rights regarding a particular a copy at the time that copy is first sold.⁶⁷ Upon that copy's sale, the owner is compensated and cannot use copyright to control its subsequent distribution.

2. *Essential step defense*

With respect to software, however, the first sale doctrine alone cannot protect consumers to the same extent that it operates with respect to books. For novels and cookbooks, the first sale doctrine allows their readers to resell the products through second-hand bookstores, and likewise allows libraries to obtain copies and lend them to patrons. But that doctrine attaches solely to copyright law's *distribution* right; it has no purchase on the coordinate *reproduction* right. In other words, the first sale doctrine affords users no ability under the law to *copy* the subject work. Yet, given that readers have no need to photocopy novels or rewrite recipes, the first sale doctrine affords them all the latitude they need.⁶⁸

But things are very different in the software world. Just to enjoy the very software copy they have acquired,⁶⁹ users need to make reproductions of it. Consider the difference—after John buys a book containing a novel, he can read and reread it a thousand times, no copying necessary; after Jane buys a DVD containing software, by contrast, she cannot use it even once on her PC without copying its contents. Software users typically need to make one permanent copy in the process of installing the software on the computer's hard disk or other persistent storage memory. When consumers use a program, a partial copy of the software is loaded from the persistent memory

67. Nonetheless, we will see below the possibility of unprecedented stress over even that conclusion. *See infra* Subsection II.B.3.

68. *Cf.* Stephen McIntyre, *Game Over for First Sale*, 29 BERKELEY TECH. L.J. 1, 11 (2014) (“The significance of the first sale doctrine can hardly be overstated. This longstanding and fundamental limitation on the public distribution right promotes commerce, safeguards competition, and props up entire industries.”) (internal citations and quotations marks omitted).

69. We use the word “acquired” here in a general sense, in order to avoid prejudicing the inquiry to come: Does a user who *acquires* diskettes containing computer software by paying \$500 on a one-time basis to the manufacturer thereby “purchase” the physical good in question, or succeed to a different legal relationship with respect to that physical good? Note that the question is *not* that user's relationship to the copyright in question, which is an intangible. *See infra* Subsection II.C.1. Instead, the first sale question arises based on the user's relationship to a physical product. Items in point could include diskettes, thumb drives, hard drives, or CD-ROMs. Properly construed, those items would even include a digital download of bits. *See* 2 NIMMER & NIMMER, *supra* note 62, at § 8.12[E].

into the computer's smaller and more agile random access memory (RAM), inasmuch as not all portions of the program code are typically needed for any given usage session.⁷⁰ The computer's central processing unit (CPU) then accesses the partial software copy in RAM and, during the operation, reproduces some smaller amount of the code in the cache memory of the CPU.⁷¹ The data resides in that cache memory only for extremely short periods, where it can be accessed and executed much faster than in RAM.⁷² During the operation of the CPU, even less data is copied into the CPU registers, comprised of a minuscule amount of memory inside the CPU.⁷³ The closer the memory is located to the CPU, where all data are processed via "current" and "no current" signals, the more accessible (faster) and smaller in size it can be, and the shorter the time it needs to be stored. These reproduction activities are dictated entirely by external functionality requirements (storage and execution efficiency), and are unrelated to particular software products or any creative expression embodied in them. Such ephemeral reproductions of small excerpts of code in cache memory and CPU registers do not typically exist long enough to meet the fixation requirement of the U.S. Copyright Act and therefore do not qualify as "copies."⁷⁴ Moreover, even if code fragments may remain in cache memory for a few seconds in atypical circumstances, they usually consist of extremely short commands bereft of copyrightability.⁷⁵ By contrast, software portions

70. In general terms, there are differences in the relative amounts of space available in a hard disc, RAM, cache, and registers. For example, Microsoft Office suite has many features that are not required in day-do-day operation. Most Word installations do not need to copy Equations Editor, and not all design templates are needed in Power Point. Thus, for execution (i.e., use), computers create "excerpts" of software, based on purely functional considerations, designed to maximize efficiency.

71. The typical label here is L1 and L2 cache.

72. See Paul Genua, *A Cache Primer*, FREESCALE SEMICONDUCTOR (2004), http://www.freescale.com/files/32bit/doc/app_note/AN2663.pdf (last visited Oct. 18, 2014); *CPU Cache*, WIKIPEDIA.ORG, http://en.wikipedia.org/wiki/CPU_cache (last visited Oct. 18, 2014).

73. A register is a small amount of very fast memory in a CPU providing quick access to commonly used values. See, e.g., *CPU Register*, WIKIPEDIA.ORG, http://en.wikipedia.org/wiki/CPU_register (last visited Oct. 18, 2014).

74. See *Cartoon Network LP, LLLP v. CSC Holdings, Inc.*, 536 F.3d 121 (2d Cir. 2008). The Act defines "copies" as "material objects, other than phonorecords, in which a work is fixed by any method now known or later developed, and from which the work can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device." 17 U.S.C. § 101 (2012).

75. See *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1156 (9th Cir. 2007). Extremely short commands either will not contain enough creative material to meet the originality requirements of the Copyright Act or will fall into carve-outs for functional subject matter under the Act; see also *supra* Subsection II.A.2. A collection and creative

reproduced in RAM often contain enough material bearing creative expression and tend to be sufficiently fixed to qualify as a “copy” eligible for protection under U.S. copyright law. Hence, RAM copies tend to implicate the reproduction right belonging to the copyright owner.⁷⁶

The first sale doctrine, being limited to the distribution right, confers no such reproduction right. The upshot, as we have just seen, is that users risk infringement liability simply by running the very software copy they have acquired. Congress redressed this situation through targeted legislation—the same 1980 amendment that conferred protection on software simultaneously added a provision allowing any lawful owner⁷⁷ of a software copy to make such additional copies as are necessary to use the acquired software.⁷⁸ The provision in question is codified as § 117 of the Copyright Act. Like § 109 (the statutory codification of the first sale doctrine), this provision applies to “the owner of a copy of a computer program.”⁷⁹

3. Reverse engineering

From a technical perspective, in order to function in combination, programs have to be interoperable, i.e., capable of exchanging and mutually using information.⁸⁰ Software manufacturers typically try to ensure that their own programs are interoperable with each other. This allows software manufacturers to market seamlessly integrated software suites, thus encouraging a broader consumer adoption of their products. As such, software manufacturers have different agendas regarding programs made by other companies and depending on the market situation. For example,

organization of short phrases as a whole can qualify as copyrightable. *See Oracle Am., Inc. v. Google Inc.*, 750 F.3d 1339, 1362 (Fed. Cir. 2014), but note that short lines of code separately copied briefly in cache to execute the code may not qualify.

76. *See Stenograph L.L.C. v. Bossard Assocs., Inc.*, 144 F.3d 96, 101–02 (D.C. Cir. 1998); *MAI Sys. Corp. v. Peak Computer, Inc.*, 991 F.2d 511, 518 (9th Cir. 1993). *See generally* Vincent J. Rocchia, *What's Fair is (Not Always) Fair on the Internet*, 29 RUTGERS L. J. 155 (1997).

77. As will be seen by the wording of the statute quoted in the next footnote, the copy that has been acquired, in order to fall under the first sale doctrine, must be *owned* by the user. Thus arises the *Vernor v. Autodesk, Inc.* construction to be explicated at length below. 621 F.3d 1102 (9th Cir. 2010).

78. “[I]t is not an infringement for the owner of a copy of a computer program to make or authorize the making of another copy or adaptation of that computer program provided: (1) that such a new copy or adaptation is created as an essential step in the utilization of the computer program in conjunction with a machine and that it is used in no other manner.” 17 U.S.C. § 117(a)(1) (2012).

79. *Id.* § 117(a). The language of the other provision actually refers more generally to “the owner of a particular copy” of a program. *Id.* 109(a) (2012).

80. *See, e.g., Interoperability*, WIKIPEDIA.ORG, <http://en.wikipedia.org/wiki/Interoperability#Software> (last visited Dec. 29, 2014).

companies with an established platform may wish, at times, to prevent interoperability with third-party software in order to protect their market share for add-on programs⁸¹ or to foster hardware sales.⁸² But new market entrants may try to achieve or promote interoperability with third-party software in order to establish their platforms or to be able to offer add-ons or substitute programs for already established and widely adopted platforms.⁸³

In the interest of optimizing the balance between exclusion and access rights, and to further innovation, courts have invoked copyright law's fair use doctrine⁸⁴ to allow intermediate copying of software code for purposes of reverse engineering and creating interoperable⁸⁵ (and even substituting⁸⁶) software products. Courts have also allowed the circumvention of technological protection measures to achieve interoperability of software-hardware combinations.⁸⁷ Those constructions underwrite copyright law's purpose of promoting creativity and innovation in the software field.

C. DIVERGENT INTERPRETATIONS OF SOFTWARE COPYRIGHT IN AND BETWEEN THE UNITED STATES AND THE EUROPEAN UNION

1. *Starting in the United States*

Given that software constitutes a form of "literary work" in the eyes of the Copyright Act,⁸⁸ and that traditional exemplars of that genre—paperback novels, fine cookbooks, and multi-volume textbooks—can be freely vended by second-hand bookstores and lent out to library patrons, it follows that software is subject to the same exploitations. The danger thereby arises that one customer buys an expensive software suite and rents it to successive

81. *See* *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992).

82. *Sony Computer Entm't, Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000), *cert. denied*, 531 U.S. 871 (2000); *Lexmark Int'l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522 (6th Cir. 2004); *Chamberlain Grp., Inc. v. Skylink Techs., Inc.*, 381 F.3d 1178 (Fed. Cir. 2004).

83. Cerulean Studios, for instance, offers a software program entitled "Trillian." Trillian is essentially a third-party user interface for various instant-messaging programs. It claims to support AOL's Instant Messenger ("AIM"), ICQ, Microsoft's MSN Messenger, Yahoo!'s Messenger, and traditional Internet Relay Chat ("IRC"). It acts as both an add-on and a full substitute for all the major instant-messaging programs. It does so without requiring the end-user to install the various instant-messaging software programs that it interfaces with. *See* CERULEAN STUDIOS, <http://www.ceruleanstudios.com/> (last visited May 31, 2013).

84. *See* 17 U.S.C. § 107 (2012).

85. *See* *Sega*, 977 F.2d at 1510.

86. *Connectix*, 203 F.3d at 596.

87. *Skylink*, 381 F.3d at 1178; *Lexmark*, 387 F.3d at 522.

88. *See* 17 U.S.C. § 101 (2012).

“library customers” for them to make their own copies. Congress took explicit note of that danger. In 1990, it amended the first sale doctrine to forbid rental of software.⁸⁹ At present, libraries are therefore forbidden from lending software, except under specified circumstances (such as those that operate at university libraries).⁹⁰

Of course, the amendment leaves software users perfectly free to sell the copies of software in their possession,⁹¹ even if they cannot rent it. Software companies have tried to avoid the implications of that state of affairs by labeling their transactions with customers as “licenses” rather than as “sales” of the physical medium.⁹² These companies have drafted shrinkwrap agreements to the effect that software copies are only licensed, never sold, and that they never transfer ownership to those physical products embodying their software. Without an authorized first sale, software copies cannot even be used on a computer without additional permission from the copyright owner.⁹³ Thus, someone in possession of a copy of software cannot use it for its intended purpose, except insofar as those rights are explicitly granted in the subject “license.”⁹⁴

At this point, we enter highly contested terrain. Dissension extends to courts, scholars, and countries. Without purporting to cover the field, some highlights deserve brief mention.⁹⁵ One point of view, adopted by some

89. The amendment in question is the Computer Software Rental Amendments Act of 1990, Pub. L. No. 101-650, § 801, 104 Stat. 5089. It is codified as a part of the first sale doctrine, as it limits the user privileges that would otherwise be as available to owners of copies of software as is available to owners of copies of novels and cookbooks, namely to rent them out.

90. See 17 U.S.C. § 109(b)(1)(A) (2012). For an analysis, see 2 NIMMER & NIMMER, *supra* note 62, at § 8.12[B][8].

91. Below, this Article will draw further distinctions between possession and ownership.

92. The question pointedly is *not* whether a license or sale of the intangible copyright has been acquired. Plainly, someone who pays \$25 for the paper and cardboard comprising a novel is only a licensee of the copyright, the same as someone who pays \$500 for set of CD-ROMs comprising a software suite is similarly only a licensee of that copyright. See *supra* note 69.

93. See *supra* Part II.B.2.

94. See *supra* Part II.B.3.

95. See Jean Braucher, *Contracting Out of Article 2 Using a “License” Label: A Strategy that Should Not Work for Software Products*, 40 LOY. L.A. L. REV. 261 (2006); Brian W. Carver, *Why License Agreements Do Not Control Copy Ownership: First Sales and Essential Copies*, 25 BERKELEY TECH. L.J. 1887 (2010); Andrew Chin, *Antitrust Analysis in Software Product Markets: A First Principles Approach*, 18 HARV. J.L. & TECH. 1 (2004); William W. Fisher III, *When Should We Permit Differential Pricing of Information?*, 55 UCLA L. REV. 1 (2007); Nancy S. Kim, *The Software Licensing Dilemma*, 2008 BYU L. REV. 1103 (2008); Michael J. Madison, *Reconstructing the Software License*, 35 LOY. U. CHI. L.J. 275 (2003); Gregory E. Maggs, *The Waning Importance of*

courts and scholars (including one of the current co-authors, who was born on U.S. shores), focuses on the particular medium embodying the software.⁹⁶ If that medium falls within the dominion of the user, then it is a sale.⁹⁷ Thus, just as someone who has paid \$15 for the hard- or paperback version of *The Castaway* can freely write in its margins, tear out chapter three, burn the whole, re-purpose it as a paperweight, or shellac it to use as a doorstop, the same considerations apply to someone who has paid \$490 for a set of CD-ROMs that embody AutoCAD software. That individual is likewise free to punch a hole in disc two, re-purpose disc one as a Frisbee, burn all three discs, or shellac the set to use as a doorstop. In *Vernor v. Autodesk, Inc.*, at the trial court level, Judge Richard Jones adopted this first viewpoint concerning AutoCAD software.⁹⁸

The contrary point of view, adopted by some courts and scholars (including the other co-author of this Article, whose birth traces back to the Continent), defers more or less to the characterization of the transaction concerning a software copy in the first agreement with the copyright owner. If the copyright owner does not agree to a sale of a software copy, then neither the first acquirer nor any downstream buyers can become an “owner” who thereby acquires rights under §§ 109 or 117 of the Copyright Act. Among followers of this viewpoint, opinions vary as to just how much deference is due the wishes of the copyright owner. Some maintain the copyright owner should be able to preclude a sale simply by stating in a form agreement that a license is intended.⁹⁹ Others believe a more substantive

Revisions to U.C.C. Article 2, 78 NOTRE DAME L. REV. 595 (2003); David A. Rice, *Copyright and Contract: Preemption After Bowers v. Baystate*, 9 ROGER WILLIAMS U. L. REV. 595 (2004); John A. Rothchild, *The Incredible Shrinking First-Sale Rule: Are Software Resale Limits Lawful?*, 57 RUTGERS L. REV. 1 (2004).

96. For a long treatment of the cases in question, with one of the current co-author's evaluation, see 2 NIMMER & NIMMER, *supra* note 62, at § 8.12[B][d][i][III].

97. Emblematic here is the following:

The Court finds that the circumstances surrounding the transaction strongly suggests that the transaction is in fact a sale rather than a license. For example, the purchaser commonly obtains a single copy of the software, with documentation, for a single price, which the purchaser pays at the time of the transaction, and which constitutes the entire payment for the “license.” The license runs for an indefinite term without provisions for renewal. In light of these indicia, many courts and commentators conclude that a “shrinkwrap license” transaction is a sale of goods rather than a license.

Softman Prods. Co., LLC v. Adobe Sys. Inc., 171 F. Supp. 2d 1075, 1085 (C.D. Cal. 2001).

98. 555 F. Supp. 2d 1164 (W.D. Wash. 2008).

99. In *MAI Systems Corp. v. Peak Computer, Inc.*, the court commented in a footnote: “Since MAI licensed its software, the Peak customers do not qualify as ‘owners’ of the

analysis is warranted to determine how “sales-like” the commercial terms of the transaction really are, regardless of the title of the agreement.¹⁰⁰ But, despite such distinctions, all varieties of this viewpoint allow the software industry to opt out of the first sale doctrine.

Defenders of this latter position focus on the economics and value-propositions of the software industry.¹⁰¹ To a large enterprise, a computer program can add immense value that is well worth the hundreds of thousands of dollars paid for a license. By contrast, a student, an educational institution, or a small business derives a much smaller economic benefit, and, concomitantly, can afford to pay much less. If the software copyright owner had to charge each user the same price to recoup development costs, the optimal price point will be somewhere in the middle. Here, the copyright owner misses out on the difference between the price charged and the price that large enterprises would be willing to pay, while also sacrificing the sales it would have made to students at a lower price. However, if copyright owners charge different prices and do not control distribution, they create an opportunity for arbitrage. For example, if a copyright owner charges students \$5 and large enterprises \$9, a student may purchase a copy and immediately sell that copy to a large company for up to \$9, pocketing the margin. And the copyright owner then misses out on differences between the educational price and large-enterprise price. Or, a large enterprise could purchase multiple licenses at favorable volume prices and sell some copies to other businesses. Here, the copyright owner misses out on an opportunity to sell single copies without discount. Given the relatively high potential for value differential with respect to computer programs and user bases, software companies have been particularly keen on controlling distribution and avoiding arbitrage.

Reverting to *Vernor v. Autodesk*, Judge Callahan on appeal vacated the judgment below.¹⁰² The Court of Appeals for the Ninth Circuit decided to

software and are not eligible for protection under § 117.” 991 F.2d 511, 518 n.5 (9th Cir. 1993).

100. Lothar Determann & Aaron Xavier Fellmeth, *Don't Judge a Sale by Its License: Software Transfers Under the First Sale Doctrine in the United States and the European Community*, 36 U.S.F. L. REV. 1, 105 (2001); see also *DSC Commc'ns Corp. v. Pulse Commc'ns, Inc.*, 170 F.3d 1354, 1360–62 (Fed. Cir. 1999).

101. See, e.g., Christian H. Nadan, *Software Licensing in the 21st Century: Are Software “Licenses” Really Sales, and How Will the Software Industry Respond?*, 32 AIPLA Q. J. 555 (2004).

102. 621 F.3d 1102 (9th Cir. 2010). Technically, this appeal relates to Judge Jones’s later opinion in this case, rather than the published opinion cited above. It is to be added that unadorned invocations of *Vernor* in the balance of this article refer to the Ninth Circuit ruling rather than to the various district court rulings in that case.

defer to the copyright owner and its *license terms*.¹⁰³ The effect, as the same appellate court expressly conceded four years earlier, is that “the first sale doctrine rarely applies in the software world because software is rarely ‘sold.’”¹⁰⁴ At this point, the prevailing view in the United States inclines towards this latter viewpoint.¹⁰⁵

Insofar as this ruling is rooted in the first sale doctrine, it applies across the board to all copyrighted works (rather than setting up a special rule for software).¹⁰⁶ Conversely, insofar as it is rooted in the essential step defense, then the matter applies solely to software.¹⁰⁷ Nonetheless, in a software copyright case (such as *Vernor*), the identical licensing terms that defeat the first sale doctrine likewise doom the essential step defense.¹⁰⁸

2. *Moving to Europe*

As we shift our focus eastward, it is necessary to note at the outset that the European legislative framework provides for more rigid versions of fair use and essential step doctrines relating to software. Under the E.U. Software Directive, a software user needs authorization from the copyright owner with respect to “the permanent or temporary reproduction of a computer program by any means and in any form, in part or in whole” including by way of “loading, displaying, running, transmission or storage of the computer program.”¹⁰⁹ But, any

person having a right to use a copy of a computer program shall be entitled, without the authorisation of the rightholder, to observe,

103. “Since the plaintiff’s license imposed ‘significant restrictions’ on the defendant’s software rights, the defendant was a licensee and was not entitled to the essential step defense.” *Id.* at 1110.

104. *Wall Data Inc. v. L.A. Cnty. Sheriff’s Dep’t*, 447 F.3d 769, 785 n.9 (9th Cir. 2006).

105. *See* *Softman Prods. Co., LLC v. Adobe Sys. Inc.*, 171 F. Supp. 2d 1075, 1085 (C.D. Cal. 2001); *DSC Commc’ns Corp.*, 170 F.3d at 1360; *Data Prods., Inc. v. Reppart*, 18 U.S.P.Q.2d 1058, 1601 (D. Kan. 1990); *ISC-Bunker Ramo Corp. v. Altech, Inc.*, 765 F. Supp. 1310, 1314 (N.D. Ill. 1990); *Davidson & Assocs., Inc. v. Internet Gateway*, 334 F. Supp. 2d 1164, 1177 (E.D. Mo. 2004); *Microsoft Corp. v. Harmony Computers & Elecs., Inc.*, 846 F. Supp. 208, 213 (E.D.N.Y. 1994); *Novell, Inc. v. Network Trade Ctr., Inc.*, 25 F. Supp. 2d 1218, 1230 (D. Utah 1997).

106. To derive the proper construction of the first sale doctrine, the court looked to past precedent ranging across all subject matters, from film prints to software itself.

107. *Vernor*, 621 F.3d at 1107. As noted above, Congress added the essential step defense of § 117 in the context of the 1980 computer amendments, applicable solely to software.

108. *Id.* (The reason is that those doctrines work in tandem. As the Ninth Circuit itself commented in this case, “both of these affirmative defenses are unavailable to those who are only licensed to use their copies of copyrighted works”).

109. Directive 2009/24/EC, *supra* note 54, art. 4(1)(a).

study or test the functioning of the program in order to determine the ideas and principles which underlie any element of the program if he does so while performing any of the acts of loading, displaying, running, transmitting or storing the program which he is entitled to do.¹¹⁰

This statutory right cannot be restricted by contract.¹¹¹ In addition to this statutory right to copy for purposes of reverse engineering, Article 5(1) of the E.U. Software Directive provides a defense similar to § 117(a) of the U.S. Copyright Act. If and to the extent reproduction is “necessary for the use of the computer program by the lawful acquirer in accordance with its intended purpose, including for error correction,” the software user does not need authorization from the copyright owner, except as otherwise agreed via contract.¹¹² Thus, if the copyright owner does not address the topic in a software license agreement, or if a secondary user is not in privity of contract, the lawful user is free to reproduce the software copy as necessary for the use of the program in accordance with its intended purpose. These statutory rights extend to lawful users and acquirers respectively, not only to owners of a copy as under the comparable feature of the U.S. Copyright Act.

The right to “communicate to the public” copyrighted works under the E.U. Copyright Directive (for example, by streaming or arranging for the download of music files over the internet) is not exhausted by a first communication.¹¹³ But a first sale of a software or music file exhausts distribution rights under the E.U. Software Directive and E.U. Copyright Directive, respectively. The first sale doctrine, as applied to software copyrights, is framed in the same way as its counterpart in the U.S. Copyright Act: “The first sale in the Community of a copy of a program by the rightholder or with his consent shall exhaust the distribution right within the Community of that copy, with the exception of the right to control further rental of the program or a copy thereof.” Notably, a sale is required and mere *acquirers* or *lawful users* of software copies are not expressly protected (unlike with respect to reverse engineering and reproduction as an essential step of using software).

110. Directive 2009/24/EC, *supra* note 54, art. 5(3).

111. Directive 2009/24/EC, *supra* note 54, art. 8.

112. Directive 2009/24/EC, *supra* note 54, arts. 5(1), (3). Contractual clauses that seek to restrict use of the software program for purposes of decompilation are invalid, *see* Directive 2009/24/EC, *supra* note 54, art. 8; Case C-406/10, SAS Inst. Inc. v. World Programming Ltd., 2012 E.C.R. I-0000, at ¶¶ 57, 58.

113. Directive 2001/29/EC, *supra* note 55, art. 3(3).

Yet, in Europe, courts have taken a different viewpoint on what kinds of transactions qualify as sales.¹¹⁴ In general, European courts give much less weight to contract terms, particularly in standard contracts imposed by software companies in the form of shrinkwrap or click-through license agreements.¹¹⁵ Moreover, E.U. law is strongly opposed to allowing companies to segment the “Common Market” in the EEA. As the E.U.’s “engine of integration,”¹¹⁶ the European Court of Justice propels the effort to strengthen European economic unification. Since copyrights are territorial and, in the thirty-one EEA states, convey national rights, copyright laws tend to obstruct—or allow companies to obstruct—rather than galvanize, borderless trade in Europe. Accordingly, the E.U. Court of Justice has, to intellectual property’s detriment, a predilection for recognizing copyright defenses and exceptions.¹¹⁷

Rooted in policy principles, on July 3, 2012, the E.U. Court of Justice held in *Oracle v. Usedsoft* that a software copyright owner may not prevent the resale of software copies that are downloaded with the copyright owner’s consent over the internet, notwithstanding the initial acquirer’s earlier agreement with the software copyright owner that the software copies are licensed only to the initial acquirer and shall not be resold.¹¹⁸ The E.U. Court embraces the view—one previously taken by German courts¹¹⁹—that any transfer of possession without a time limit for a lump sum fee constitutes a sale and triggers the first sale doctrine.¹²⁰

The E.U. Court also expands this view to apply to software downloads and indicates that someone who acquires a software copy lawfully, i.e., from the copyright owner, with the copyright owner’s consent or from a secondary

114. See Determann & Fellmeth, *supra* note 100; see also Lothar Determann, *Importing Software and Copyright Law*, 30 COMPUTER & INTERNET LAW. 1, 4 (2013).

115. See Lothar Determann, *Notice, Assent Rules for Contract Changes After Douglas v. U.S. District Court*, 12 BNA ELECT. COM. & L. REP. 1 (2007); James R. Maxeiner, *Standard-Terms Contracting in the Global Electronic Age: European Alternatives*, 28 YALE J. INT’L L. 109 (2003); see also Jessica Gallegos, *A New Role for Tortious Interference in the Digital Age: A Model to Enforce End User License Agreements*, 38 FLA. ST. U. L. REV. 411 (2011); Mark A. Lemley, *Intellectual Property and Shrinkwrap Licenses*, 68 S. CALIF. L. REV. 1239, 1246, 1261, 1267–84 (1995); Margaret Jane Radin, *Reconsidering Boilerplate: Confronting Normative and Democratic Degradation*, 40 CAP. U. L. REV. 617 (2012).

116. See, e.g., Matthew Gabel, *The European Court of Justice as an Engine of Economic Integration: Reconsidering Evidence that the ECJ has Expanded Economic Exchange in Europe* (Aug. 5, 2009), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1444500.

117. *Importing Software and Copyright Law*, *supra* note 114, at 4–5.

118. Case C-128/11, *UsedSoft GmbH v. Oracle Int’l Corp.*, 2012 E.C.R. I-0000, 2012 O.J. (C 287) 16, at ¶¶ 44–46; see Determann & Batchelor, *supra* note 65.

119. See Determann & Fellmeth, *supra* note 100, at 92–101.

120. See *Importing Software and Copyright Law*, *supra* note 114, at 5.

distributor after exhaustion kicks in, may make and sell an additional copy so long as the original software copy is made “unusable,”¹²¹ a view that has recently been rejected in the United States with respect to digital music files.¹²² Consequently, downloaded software copies can be resold much more easily in Europe, because they can be freely separated from media or devices where they are originally installed.

Going even further, and on unclear legal grounds, the E.U. Court indicates that, after copyright exhaustion kicks in, secondary purchasers may transfer licenses relating to software copies that are transferred in sales-like transactions.¹²³ While the first sale doctrine does not itself address the transfer of license agreements, it seems that the E.U. Court of Justice expands the scope of the first sale doctrine from copyright to contracts law based on mere policy considerations, to give the doctrine force. Finally, the E.U. Court of Justice ruled that any contractual agreements to the contrary would not be enforceable, regardless of party sophistication or equal bargaining strength.

In short order, German courts blessed the resale of software copies that were first licensed to educational institutions (subject to heavy discounts and restrictive licenses) and subsequently sold to UsedSoft for the purposes of further resale and margin arbitrage.¹²⁴

Emblematic of the turmoil in this field is that U.S. courts have moved in one direction, and European courts in the opposite. Scholars remain divided; one need only add that the American co-author of this Article sides with the governing view in Europe, whereas the European co-author sides with the precedential view articulated in the United States.

3. *Cross-breezes: exceptions in United States and E.U. to the paradigms set forth above*

Life is never static; the winds that have blown easterly may switch west on a moment’s notice. The cases discussed above, in like measure, established rules that have been buffeted and subverted by more recent cases

121. Case C-128/11, *UsedSoft GmbH v. Oracle Int’l Corp.*, 2012 E.C.R. I-0000, 2012 O.J. (C 287) 16, at ¶ 70.

122. *See* *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640, 650 (S.D.N.Y. 2013) (stating that, for the purposes of violating the reproduction right, it is irrelevant whether the original phonorecord no longer exists after making a copy).

123. Case C-128/11, *UsedSoft GmbH v. Oracle Int’l Corp.*, 2012 E.C.R. I-0000, 2012 O.J. (C 287) 16, at ¶¶ 67, 84–85.

124. *See, e.g.*, *OLG Frankfurt am Main* (German Court of Appeals), Nov. 6, 2012, Az. 11-U68-11 (*Adobe v. Usedsoft*).

in both the New World and the Old. A brief foray into subsequent jurisprudence is therefore indicated.

a) Returning Stateside

The latest hurricane from the United States arose in its construction of the first sale doctrine of U.S. copyright law in the context of literary works (albeit in the form of textbooks, rather than computer software). Until recently, it was widely assumed that a sale abroad of copies made outside the U.S. would not exhaust distribution rights under U.S. copyright law, because copyrights are territorial.¹²⁵ In Europe, for example, the E.U. Court of Justice has held that sales in the United States or elsewhere outside the EEA do not exhaust intellectual property rights within the EEA.¹²⁶ So the mirror result was contemplated under U.S. law, in accordance with the very first opinion to address that thorny issue.¹²⁷

But reality diverged from expectations. In *Kirtsaeng v. John H. Wiley & Sons*,¹²⁸ the U.S. Supreme Court ruled that copies lawfully made and first sold abroad could be imported into the United States and resold without the U.S. copyright owner's consent. Supap Kirtsaeng had imported books from Thailand and asserted the first sale doctrine as a defense when a U.S. copyright owner, John Wiley & Sons, Inc., sued him for copyright infringement.¹²⁹ The Second Circuit had rejected the defense based on territoriality considerations, holding the first sale defense inapplicable insofar as the copies were not made in the United States, meaning that no authorized first sale had occurred in the United States.¹³⁰ But in a majority opinion written by Justice Breyer, the Supreme Court reversed, holding that a first sale outside the U.S. qualifies for the defense.¹³¹

125. "Section 109(a), properly read, affords Kirtsaeng no defense against Wiley's claim of copyright infringement. The Copyright Act, it has been observed time and again, does not apply extraterritorially." *Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1376 (2013) (Ginsburg, J., dissenting).

126. Case C-479/04, *Laserdisken ApS v. Kulturministeriet*, 2006 E.C.R. I-8089, 30, ¶ 24; see also Case C-355/96, *Silhouette Int'l Schmied GmbH & Co. KG v. Hartlauer Handelsgesellschaft mbH*, 1998 E.C.R. I-4799 (regarding exhaustion under trademark law).

127. *Columbia Broad. Sys., Inc. v. Scorpio Music Distribs., Inc.*, 569 F. Supp. 47 (E.D. Pa. 1983), *aff'd mem.*, 738 F.2d 424 (3d Cir. 1984). Notwithstanding that 1983 opinion, the Third Circuit later came to basically the opposite conclusion. See *Sebastian Int'l, Inc. v. Consumer Contacts (PTY) Ltd.*, 847 F.2d 1093, 1099 (3d Cir. 1988).

128. *Kirtsaeng*, 133 S. Ct. 1351.

129. *Id.* at 1357–58.

130. *John Wiley & Sons, Inc. v. Kirtsaeng*, 654 F.3d 210 (2d Cir. 2011), *rev'd*, 133 S. Ct. 1351 (2013).

131. *Kirtsaeng*, 133 S. Ct. at 1371. Previously, the Court had ruled that U.S.-made copies that had been sold abroad with the U.S. copyright owner's authorization could be imported

Although *Kirtsaeng* dealt with books, the Supreme Court examined a variety of factors favoring international exhaustion of the copyright owner's exclusive right of distribution, including the implications for software.¹³² Neither the parties nor the justices questioned whether Kirtsaeng owned the copies that he sold in the United States. Book publishers have not historically tried to draw a distinction between ownership to the physical book and a "license" of that physical item. Unlike software companies, book publishers have been content to "sell" copies of their works and have not required distributors and end users to accept "license agreements" that purportedly allow only certain uses and concomitantly forbid transfers.¹³³

The *Kirtsaeng* majority found comfort in the fact that its interpretation of the first sale doctrine in current U.S. copyright law is aligned with the common law version of the first sale doctrine relating to the transfer of ownership to chattels that comes with an "impeccable historic pedigree."¹³⁴ Citing seventeenth century opinions relating to property, the Supreme Court notes, "[t]he 'first sale' doctrine also frees courts from the administrative burden of trying to enforce restrictions upon difficult-to-trace, readily movable goods."¹³⁵

Of course, that sensibility is exactly contrary to the Ninth Circuit's ruling in *Vernor v. Autodesk*, which is entirely premised on enforcing restrictions upon readily movable goods. Moreover, *Kirtsaeng* also challenges the twin rationales on which *Vernor* has been defended—its blanket obeisance to terms embodied by copyright owners onto copies of their works, and its facilitation of price discrimination.

The *Kirtsaeng* majority viewed its ruling as vindication of the Court's own earlier ruling in *Bobbs-Merrill Co. v. Straus*, stating that "copyright laws were not 'intended to create a right which would permit the holder of the

against the U.S. copyright owner's will. See *Quality King Distribs., Inc. v. L'Anza Research Int'l, Inc.*, 523 U.S. 135 (1998).

132. "[A]utomobiles, microwaves, calculators, mobile phones, tablets, and personal computers' contain copyrightable software programs or packaging. . . . Many of these items are made abroad A geographical interpretation would prevent the resale of, say, a car, without the permission of the holder of each copyright on each piece of copyrighted automobile software." *Kirtsaeng*, 133 S. Ct. at 1365.

133. Nonetheless, they have relied on geographical exclusivity on those sold copies. See, for example, the notice that the copyright owner in *Kirtsaeng* included in its books: "Copyright 2008 John Wiley & Sons (Asia) Pte Ltd. All rights reserved. This book is authorized for sale in Europe, Asia, Africa, and the Middle East only and may be not exported out of these territories." J. WALKER, *FUNDAMENTALS OF PHYSICS* vi (Wiley Int'l Student ed., 8th ed. 2008).

134. *Kirtsaeng*, 133 S. Ct. at 1353.

135. *Id.* at 1363.

copyright to fasten, by notice in a book . . . a restriction upon the subsequent alienation of the subject-matter of copyright after the owner had parted with the title to one who had acquired full dominion over it.”¹³⁶ In that 1908 case, the copyright owner had purported to impose its own license terms: “The price of this book at retail is one dollar net. No dealer is licensed to sell it at a less price, and a sale at a less price will be treated as an infringement of the copyright.”¹³⁷ Of course, the Court’s refusal to bow to those “license” terms is precisely what gave birth to the first sale doctrine that, over a century later, *Kirtsaeng* continued to vindicate.

Moving to price discrimination,¹³⁸ the Court’s previous foray into this domain had unanimously, in a single sentence, rejected that feature as a basis for its decision.¹³⁹ This time, although the dissent would have upheld price discrimination,¹⁴⁰ the six justices in the majority, conceding that their interpretation might make price discrimination impossible, nonetheless explicitly disclaimed that consideration as legally cognizable for purposes of copyright doctrine.¹⁴¹

136. *Id.* at 1367 (quoting *Bobbs-Merrill Co. v. Straus*, 210 U.S. 339, 349–50 (1908)) Most of that same quotation, in turn, was earlier articulated by the unanimous Court in its earlier opinion. *See Quality King Distrib., Inc. v. L’Anza Research Int’l, Inc.*, 523 U.S. 135, 141 (1998).

137. *Bobbs-Merrill*, 210 U.S. at 341.

138. Reverting to *Vernor v. Autodesk, Inc.*, the plaintiff copyright owner urged that same argument. 621 F.3d 1102, 1114–15 (9th Cir. 2010). “Autodesk contends that this (1) allows for tiered pricing for different software markets, such as reduced pricing for students or educational institutions; (2) increases software companies’ sales; (3) lowers prices for all consumers by spreading costs among a large number of purchasers; and (4) reduces the incidence of piracy by allowing copyright owners to bring infringement actions against unauthorized resellers.” The Ninth Circuit allowed that there are “serious contentions on both sides” of the policy argument. *Id.* at 1115.

139. *Quality King Distrib.*, 523 U.S. at 153 (“[W]hether or not we think it would be wise policy to provide statutory protection for such price discrimination is not a matter that is relevant to our duty to interpret the text of the Copyright Act.”).

140. “Because economic conditions and demand for particular goods vary across the globe, copyright owners have a financial incentive to charge different prices for copies of their works in different geographic regions. Their ability to engage in such price discrimination, however, is undermined if arbitrageurs are permitted to import copies from low-price regions and sell them in high-price regions. The question in this case is whether the unauthorized importation of foreign-made copies constitutes copyright infringement under U.S. law.” *Kirtsaeng*, 133 S. Ct. at 1374 (Ginsburg, J., dissenting).

141. An excerpt reads as follows:

Wiley and the dissent claim that a nongeographical interpretation will make it difficult, perhaps impossible, for publishers (and other copyright holders) to divide foreign and domestic markets. We concede that is so. A publisher may find it more difficult to charge different prices for the same book in different geographic markets. But we do not see how these facts

The stark contrast between *Vernor* and *Kirtsaeng* comes into high relief when one reflects that a full-throttled application of the former could have obviated any need for the latter even to arise. Taken at face value, the Ninth Circuit's pronouncement that its ruling applies across the board to all copyrightable works¹⁴² opens the door for the publishing industry to adopt the same expedients used by software manufacturers. At the dawn of software commercialization, purveyors had sold products for a fixed sum, with the tangible medium remaining perpetually with the purchaser.¹⁴³ At a later point in time, the same companies began to unilaterally style the transaction as a "license" of products for a fixed sum, with the tangible medium remaining perpetually with the "licensee"—but subject to a "license term" against its further distribution. The Ninth Circuit's vindication of that stratagem as a general matter of copyright law thus furnishes a clear road-map to John Wiley & Sons and its fellow purveyors of textbooks: whereas they used to sell a given volume for \$45, now they need only wrap it in cellophane (covering a printed form limiting user rights) and pronounce the resulting transaction at the same \$45 price point a "license." In that manner, Wiley et al. not only would be able to forestall importation of the subject books, but also would likewise be able to prevent the books from being sold at second-hand bookstores, or even lent from libraries. In this way, the Ninth Circuit's previous recognition that "the first sale doctrine rarely applies"¹⁴⁴ would be generalized from "the software world" to the entire panoply of copyrightable expression.

To be sure, it is easy to imagine that a future Ninth Circuit panel would reaffirm application of *Vernor* to the world of books, even after *Kirtsaeng*. Its

help Wiley, for we can find no basic principle of copyright law that suggests that publishers are especially entitled to such rights.

. . . . [T]he Constitution's language nowhere suggests that its limited exclusive right should include a right to divide markets or a concomitant right to charge different purchasers different prices for the same book, say to increase or to maximize gain. Neither, to our knowledge, did any Founder make any such suggestion. We have found no precedent suggesting a legal preference for interpretations of copyright statutes that would provide for market divisions.

To the contrary, Congress enacted a copyright law that (through the "first sale" doctrine) limits copyright holders' ability to divide domestic markets. And that limitation is consistent with antitrust laws that ordinarily forbid market divisions.

Id. at 1370–71 (citations omitted).

142. See *supra* Subsection II.C.3.a)

143. See generally Leonard J. Kennedy & Lori A. Zallaps, *If it Ain't Broke . . . The FCC and Internet Regulation*, 7 COMMLAW CONSPECTUS 17 (1999).

144. See *supra* text accompanying note 104.

logic would be: *We are bound to give expansive reach to the first sale doctrine by governing Supreme Court authority; but plaintiffs Random House and Harcourt Brace at bar are no longer engaging in "sale" of their works, so that doctrine has no purchase anymore; instead, the subject books should be viewed as "licensed" only.* Yet it seems equally possible that a future Supreme Court, confronted with the issue, would reject that extension of *Vernor's* construction. We return to those considerations presently.¹⁴⁵

b) Final European Jaunt

By the same token, finality is evanescent across the Atlantic as well. In particular, brief mention should be made of two recent decisions at the trial court level, one limiting *Oracle's* ruling to software¹⁴⁶ and the other potentially expanding the effect of the decision exponentially.

On March 5, 2013, a German consumer protection watchdog lost a lawsuit against the operator of a website that was selling non-transferable licenses to electronic books and music files. The watchdog had complained that the contractual resale prohibitions were invalid under *Oracle v. UsedSoft*¹⁴⁷ and therefore constituted unfair business practices, because they confused consumers about their rights. The German court disagreed, holding that the contractual resale prohibitions were not unfair. The court noted that *Oracle v. UsedSoft* was based on the E.U. Software Directive, relating to exhaustion of distribution rights, whereas the download of electronic books and music files constituted a "communication to the public" under the E.U. Copyright Directive, which is expressly not subject to exhaustion.¹⁴⁸ Thus, the German court came out against a general digital first sale doctrine, just as a U.S. court did a few weeks later in *Capitol Records v. Redigi*.¹⁴⁹

On April 17, 2013, the pendulum swung in the opposite direction, this time against software copyright owners. The U.K. Supreme Court held that loading a webpage does not trench on the reproduction rights of a copyright

145. See *infra* Subsection II.C.4.

146. As noted above, the Ninth Circuit in *Vernor v. Autodesk, Inc.*, announced that its ruling applied across the board and was not limited to the software realm. 621 F.3d 1102 (9th Cir. 2010).

147. Case C-128/11, *UsedSoft GmbH v. Oracle Int'l Corp.*, 2012 E.C.R. I-0000, 2012 O.J. (C 287) 16, discussed *supra* Part II.C.2.

148. Bielefeld Regional Court decision of March 5, 2013, see press release of the plaintiff, a German consumer protection center, at <http://www.vzbv.de/11533.htm> and the court decision at http://www.vzbv.de/cps/rde/xbcr/vzbv/LG_Bielefeld-E_Books-2013-03-05.pdf (both visited Nov. 23, 2014).

149. *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640 (S.D.N.Y. 2013).

owner.¹⁵⁰ Accordingly, a user who accesses a remotely hosted software copy via the internet and a locally cached copy of the software GUI does not require the software copyright owner's permission. Paired with the holding in *Oracle v. UsedSoft*, a software user may be permitted not only to resell a software copy to one secondary buyer, but to also make it available to thousands of software users as part of a cloud offering, if the same result can be technically structured in a manner that avoids additional copying of the underlying code. Again, we return to those considerations presently.¹⁵¹

4. *Departure from the statute's goals?*

The above rundown reveals severe climatic changes in recent decades. When adopted, software was slotted into the realm of "literary works" protected by copyright, so as to afford rights to proprietors while simultaneously reserving salient privileges to users. The safeguards of pre-existing copyright law applied automatically—free alienability of tangible manifestations via the first sale doctrine and application of copyright's all-purpose defense, the fair use doctrine. In addition, the 1980 law that recognized software protection concomitantly relaxed the reproduction strictures applicable to this new domain, by virtue of the tailored addition of § 117 to the statute.

Yet, as we have seen, *Vernor* and similar cases recognize plenary protection for software copyrights on the proprietors' side of the ledger, without any of the corresponding safeguards on the users' side. In particular, that Ninth Circuit case overtly denies application of the first sale doctrine, through the simple expedient of replacing "for sale" with "for license" in a single-pay transaction through which possession of tangible media changes hands forever. It also explicitly rejects application of the essential step defense, which is likewise a privilege belonging only to the "owner of a copy" of the software in question.¹⁵² And in that context, it even approves the elimination, through that "license," of the user's ability to reverse engineer the software in question¹⁵³—a privilege that the fair use doctrine itself grants,

150. *Pub. Relations Consultants Ass'n Ltd. v. Newspaper Licensing Agency Ltd.*, [2013] UKSC 18.

151. *See infra* Part III.

152. Not only does *Vernor* extinguish privileges under the first sale doctrine, but its effect is equally devastating to users' coordinate privileges under the essential step doctrine. 621 F.3d 1102. *See supra* note 107.

153. As recited in the Ninth Circuit opinion, one provision in Autodesk's license stated, "YOU MAY NOT: (1) modify, translate, reverse-engineer, decompile, or disassemble the Software." *Vernor*, 621 F.3d at 1104. The Ninth Circuit expressly cited that limitation among the "use restrictions" that caused it to "conclude that [Autodesk's] customers are licensees of their copies of [the software in question] rather than owners." *Id.* at 1111–12.

according to antecedent Ninth Circuit authority.¹⁵⁴ In short, *Vernor*'s tally is: Copyright Owners 4, Users 0.¹⁵⁵

As we have further seen, there is no reason for John Wiley & Sons and others similarly situated to remain aloof from *Vernor*'s possibilities. Having lost *Kirtsaeng* at the Supreme Court level, those book publishers need only invest in a shrinkwrap machine to enclose all future inventory in a form contract proclaiming itself a "license." Not only can that expedient serve to bar importation of unwanted volumes, it can even ratchet up the publishers' rights to prevent loss of revenue occasioned by the existence of second-hand bookstores within the United States.¹⁵⁶ The magical cellophane can even keep those volumes off of library shelves,¹⁵⁷ so that the publishers would not have to make do with a single sale resulting in dozens of library patrons reading the same copy over time.¹⁵⁸ In all those ways, they may now snatch victory from the jaws of *Kirtsaeng*'s defeat.

Although book publishers would discover their newfound rights extinguishing users' privileges limited to the distribution right, their colleagues in the software business could push matters much further. Once it is established that a user in possession of a given physical product embodying

154. *See supra* Subsection II.A.3.

155. As set forth above, the metrics in place are: (1) copyright protection, (2) user's ability to resell, (3) user's ability to invoke § 117, and (4) user's ability to invoke fair use.

156. To be sure, those bookstores could continue to sell old inventory. But, once the publishers wised up to labeling the distribution of their products a "license" rather than a "sale," future resale of the volumes in question would be rendered *verboten*.

157. Copyright owners have always lacked a library lending right under United States copyright law, of the sort that Canadian and Dutch law, for example, grant them. *See* 2 NIMMER & NIMMER, *supra* note 62, § 8.12[B][1][a] n.27 (citing *International Copyright Law and Practice*).

158. Those publishers need not stop even there. As set forth in the write-up by one of the current co-authors, a broad construction of *Vernor* frees book publishers to embody "use restrictions" on readers of affected volume from undertaking any of the following prohibited activities:

- to read Chapter 5 before Chapter 4;
- to read Chapter 6 more than three times;
- to be seen reading the book in saloons or Bohemian redoubts;
- to read multiple paragraphs aloud to a friend or relative;
- to write a review of the book, absent pre-screening with the publisher, which has the unilateral right to veto any commentary not to its liking.

2 NIMMER & NIMMER, *supra* note 62, § 8.12[B][1][d][i][III]. The third item in the list derives from a parallel use restriction validated by the same panel that heard *Vernor*, in a companion copyright case. *See* MDY Indus., LLC v. Blizzard Entm't, Inc., 629 F.3d 928, 938–39 (9th Cir. 2010) ("The game must be used only for non-commercial entertainment purposes and may not be used in cyber cafes and computer gaming centers without Blizzard's permission.").

software is not the “owner” of that tangible object but rather only its “licensee,” then not only are distribution rights lost under the first sale doctrine, but so likewise are reproduction rights under the essential step defense.¹⁵⁹ Accordingly, a software purveyor may “license” copies of AutoCAD for \$490 per copy, and then turn around and sue the user for infringement the first time it dares to actually run the program by copying it into RAM. Granted, most extant End User License Agreements allow such running in RAM (in consonance with the dictates of § 117 itself).¹⁶⁰ But nothing forces copyright owners to be so magnanimous. In other words, those proprietors may provide explicitly that the copy that has been “licensed” for \$490 may never be run in RAM, absent additional conditions that the license imposes.¹⁶¹ Absent compliance, the subject reproduction through first-time running of the software in RAM becomes unlicensed, and hence infringing under *Vernor*’s logic.¹⁶²

If squarely presented with the issues just joined, it is conceivable that the U.S. Supreme Court would affirm the application of *Vernor* to all the above circumstances. It might reason that its decision in *Kirtsaeng* only construed the first sale doctrine, and that copyright owners remain free to fashion their own affairs by eliminating that doctrine entirely (along with the essential step

159. It should be recalled that the difference between computer programs and other species of literary works is such that “users risk infringement liability simply by running the very software copy they have acquired,” a situation that Congress addressed “by targeted legislation” adding the “essential step” defense. *See* discussion *infra* Subsection II.B.2

160. To the extent that the EULAs are silent, an argument could arise that the doctrine of *implied license* privileges the user who has paid \$500 for a “licensed copy” of AutoCAD to run it in RAM. *See* 3 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 10.03[A][7] (2013). But the copyright owner can effectively extinguish any defense rooted in such an *implied* license by simply including an *explicit* term in the license, forbidding that conduct unless specifically authorized along the lines to be illustrated momentarily.

161. Those terms could be anything that the copyright owner cares to embody in the EULA, such as:

prior to loading software into RAM, the user must make a phone call to the software proprietor’s headquarters to specify the exact location of the machine in question, followed by a subsequent phone call indicating that the session had ended;

prior to loading software into RAM, the user must pay an additional \$500;
or

prior to loading software into RAM, the user must sign onto Facebook and “like” all items specified in the URL supplied by the copyright proprietor.

162. “The Copyright Act confers [the essential step] defense only on owners of software copies. *See* 17 U.S.C. § 117 (2012). In contrast, a licensee’s right to use the software, including the right to copy the software into RAM, is conferred by the terms of its license agreement.” *Vernor v. Autodesk, Inc.*, 621 F.3d 1102, 1112 n.13 (9th Cir. 2010).

defense for software). If those proprietors simply shrinkwrap their goods with a form contract, the intended defenses of copyright law are set at naught. Indeed, even the cellophane expedient itself could be viewed as superfluous, with the license terms accomplishing their intended objective simply by virtue of being clearly delineated on the book's title page.

That hypothetical outcome certainly represents one possibility. But its contrary is more easily envisioned. After all, the Supreme Court's own seminal 1908 decision, later codified as the first sale doctrine, rejected the cogency of precisely such a purported "license" limitation on the book's title page.¹⁶³ Viewed through that prism, it is straightforward to imagine the current Court rejecting the publishers' postulated end-run around *Kirtsaeng*.¹⁶⁴

In short, the law is in flux. The marriage of copyright law to software protection has reached a state in which unforeseen developments have rendered the nuptial conditions suspect. Had Congress known in 1980 that software publishers would be able to unilaterally ratchet up their rights—such that they would continue to enjoy complete copyright protection over their handiwork, untrammelled by the expected limitations of the first sale doctrine, essential step defense, and the quintessential matter of fair use¹⁶⁵—it is unclear whether Congress would have even gone forward to enact the amendment. These considerations suffice to warrant another look at the legal landscape. But the situation is actually even starker. To appreciate why, we must focus our attention on the cloud.

163. In *Bobbs-Merrill Co. v. Straus*, the legend read as follows: "The price of this book at retail is one dollar net. No dealer is licensed to sell it at a less price, and a sale at a less price will be treated as an infringement of the copyright." 210 U.S. 339, 341 (1908).

164. In *Vernor*, the Ninth Circuit postulated that "the Supreme Court in *Bobbs-Merrill* made explicit that its decision did not address the use of restrictions to create a license." 621 F.3d at 1114. In fact, by contrast, the Supreme Court overtly stated in that earlier case that the publisher in that case had limited the rights that it had "licensed." *Bobbs-Merrill*, 210 U.S. at 341. It then went on to state:

The precise question, therefore, in this case is, does the sole right to vend . . . secure to the owner of the copyright the right, after a sale of the book to a purchaser, to restrict future sales of the book at retail, to the right to sell it at a certain price per copy, because of a notice in the book that a sale at a different price will be treated as an infringement, which notice has been brought home to one undertaking to sell for less than the named sum? We do not think the statute can be given such a construction

Id. at 350.

165. "Although the traditional approach is to view 'fair use' as an affirmative defense, this writer, speaking only for himself, is of the opinion that it is better viewed as a right granted by the Copyright Act of 1976." *Bateman v. Mnemonics, Inc.*, 79 F.3d 1532, 1542 n.22 (11th Cir. 1996).

III. ECONOMICS AND TECHNOLOGY OF THE CLOUD

A. SOFTWARE COMMERCIALIZATION MODELS

Software developers have always had numerous vehicles to exploit their innovations, including the following:¹⁶⁶

- *internal use*: keep the software secret and use it internally as a competitive advantage for other business activities, e.g., manufacturing, financial services, administration, business planning, marketing, or product development;
- *contract development*: sell development services and work product, including title to the intellectual property rights to the software;
- *distribution of copies*: sell, lease, or give away copies of software (possibly subject to license agreements) to enterprises or consumers for a fee, to collect personal data, to establish a platform, or to generate goodwill in the open source community;
- *distribution of hardware-software combinations*: sell or lease products with pre-installed software, e.g., personal computers, enterprise servers, laptops, smart phones, etc.

In the first business model (internal use), the software developer does not make the software available at all. In the three other models, the developer transfers possession of software copies to users. Between these all-or-nothing cases regarding transfer of possession exists another group of business models, through which a software company retains physical possession of the software copies (and the hardware on which the software runs) and enables users to remotely access and use the software functionality. Such business models have been on the rise over the last fifteen years, under such names as application service providing (“ASP”), infrastructure-as-a-service (“IaaS”), platform-as-a-service (“PaaS”), and software-as-a-service (“SaaS”).¹⁶⁷ More generally, they are known as “cloud” models,¹⁶⁸ which is a term used throughout the remainder of this Article as a collective reference for service-based software exploitation models.

Economically, cloud models feature a number of aspects traditionally associated with services and leases, but without meeting all legal elements of

166. See generally Lothar Determann, *What Happens in the Cloud: Software as a Service and Copyrights*, 29 BERKELEY TECH. L.J. 1095 (2014).

167. See Ben Kepes, *Understanding the Cloud Computing Stack: SaaS, PaaS, IaaS*, RACKSPACE, http://www.rackspace.com/knowledge_center/whitepaper/understanding-the-cloud-computing-stack-saas-paas-iaas/ (last visited Oct. 9, 2014).

168. Lothar Determann, *Data Privacy in the Cloud—Myths and Facts*, 121 PRIVACY L. & BUS. 17 (2013).

these models.¹⁶⁹ Similarly to both leases and services, cloud software users pay recurring fees and receive benefits during an agreed-upon, limited term. Unlike in traditional lease arrangements, however, the cloud user does not receive physical possession of software copies. The user is not responsible for maintaining the cloud offering (like a lessee typically would be in the context of a lease). The cloud user relies on the cloud provider with respect to software availability, operability, and data security. Unlike in the lease context, the cloud user does not return anything to the cloud provider at the end of the agreed term. Instead, the cloud provider returns the user's data.

And similarly to services, cloud users do not have to install, maintain, repair, or update hardware or software, or deal with interoperability. Unlike in traditional service arrangements, however, (e.g., word processing or outsourced accounting services), cloud users operate the software and create (and own *ab initio*) the work product generated with the cloud offering (e.g., Word documents, PowerPoint slides, Excel spreadsheets).

Cloud models do not resemble sales from an economic perspective. The cloud provider does not transfer possession to anything, the arrangement is not perpetual, and the user pays a recurring fee, rather than a one-time purchase price. Cloud models, like the internal use model, do not resemble sales. However, cloud models do not resemble the internal use model either. For example, the cloud provider does not use the software to support other business activities (such as manufacturing or other services). The software functionality *is* the service.

What advantages do cloud models have over the more traditional models summarized above? Software users obtain access to state-of-the-art software functionality without the burden of dealing with infrastructure, hardware, updates, maintenance, etc., which are not core competencies of many organizations. Also, change management (updates, upgrades, bug fixes) can be less burdensome for users. Users typically do not have to pay a large upfront sum in subscription models. Additionally, the overall costs of software and computing resources over time can be reduced as cloud providers bundle purchasing power and reduce redundancies by deploying computing capacity dynamically (i.e., only when needed by a particular user). To enjoy these benefits, more software users are willing to hand over their data to cloud providers, even though data security and third-party access to

169. See Ted Alford & Gwen Morton, *The Economics of Cloud Computing: Addressing the Benefits of Infrastructure in the Cloud*, BOOZ ALLEN HAMILTON (2010), available at <http://www.boozallen.com/media/file/Economics-of-Cloud-Computing.pdf> (last visited October 9, 2014); *The Economics of the Cloud*, <http://news.microsoft.com/download/archived/presskits/cloud/docs/the-economics-of-the-cloud.pdf> (last visited Oct. 9, 2014).

data remain major concerns to some organizations. These concerns may be especially prevalent in less traditional business models.¹⁷⁰

B. EXAMPLES OF CLOUD OFFERINGS

Cloud providers offer different commercial terms and various technological environments, for example, with respect to access (e.g., via third-party web browsers or proprietary thin client software downloaded on customers' computers), hardware configuration (e.g., one server for all customers hosted in multi-tenant data centers or one dedicated server and facility per customer) and software architecture (e.g., one copy in RAM serves multiple customers or each customer accessing the remote offering causes a separate RAM copy to be created). Offerings depend on the provider's business model, on whether the provider develops its own software or largely uses programs made by other companies, and on other factors.¹⁷¹ A given cloud provider may offer third-party software application products (such as Microsoft Word, Microsoft PowerPoint, or Adobe Acrobat) to enterprise and consumer customers.¹⁷² Another may host third-party computer games or components thereof.¹⁷³ Yet other providers develop and host their own enterprise applications for customer relationship management, human resources systems, or enterprise resource planning.¹⁷⁴

Our analysis herein refers to a simplified, typical, technical scenario: the cloud provider acquires software copies by developing them internally or procuring them from outside sources. Then, the cloud provider creates the cloud offering by combining application programs with operating system software, drivers, and programs that facilitate remote access. The cloud provider installs the software combination on server hard disks in secure locations with connections to the internet. When the cloud provider turns the cloud service on, it uploads a software copy into the RAM of one or more servers for its customers to use.

Customers enter into an agreement with the cloud provider, pay a recurring fee, and receive access credentials (user IDs and passwords).

170. Lothar Determann, *Data Privacy in the Cloud: A Dozen Myths and Facts*, 28 *COMPUTER & INTERNET LAW* 11, 1 (2011).

171. Such as software functionality, industries, targeted user groups (enterprise, consumers, prosumers, etc.).

172. *See, e.g.*, ONLIVE, <http://www.desktop.onlive.com/> (last visited Dec. 13, 2014).

173. *See generally* MDY v. Blizzard Entm't, Inc., 629 F.3d 928 (9th Cir. 2010).

174. *See, e.g.*, SALESFORCE.COM, INC., <http://www.salesforce.com/> (last visited Dec. 13, 2014); SUGARCRM, <http://www.sugarcrm.com> (last visited Oct. 13, 2014); WORKDAY, INC., <http://www.workday.com/> (last visited Dec. 13, 2014); NETSUITE INC., <http://www.netsuite.com/> (last visited Dec. 13, 2014).

Customers can then access the cloud offering with general-purpose web browsers (e.g., Firefox, Chrome, Safari, or Internet Explorer). The work product that the customer creates with the cloud offering consists of data (e.g., in the form of a PowerPoint slide deck or Word document) stored on the cloud provider's server. The customer can view the work product via the GUI that is reproduced via the web browser on the remote computer. If the customer downloads (i.e., copies) the work product to its remote computer, the cloud provider's server may deliver the work product in files that contain standard file format specifications to enable the user to process the files on remote computers. If the cloud provider's offering includes objects for inclusion into work product (e.g., clipart for PowerPoint slides), customers can view or download copies of such objects, too.

C. MULTIPLE CLOUD USERS

Cloud service providers can configure their software such that one RAM copy can simultaneously serve multiple users. In a multi-tenant, multi-threaded setup, dozens or even thousands of users can use the same RAM copy without the need to create additional copies of the application software that provides the program functionality on the remote users' computers. Does that activity implicate the copyright owner's rights? We begin with the core reproduction right,¹⁷⁵ and then radiate outward to consider the other rights belonging to the proprietor.¹⁷⁶

Counting copies for the purposes of copyright law analysis produces the following results: a company that offers software as a service via the cloud typically must create two copies. Included are one permanent copy of the underlying code on its server and one RAM copy in the working memory. In that regard, the cloud scenario does not involve more copies than a traditional desktop scenario.¹⁷⁷ However, in the cloud context, the one RAM copy can be accessed by multiple users—very much unlike the desktop scenario.

In the usage phase, once the cloud solution is up and running, users make access requests from remote computers. Each time the software is executed, numerous fractional excerpts of the RAM copy are reproduced in cache memory spaces and the CPU of the cloud provider's server. As in the desktop scenario, such fractional excerpts of the RAM copy do not implicate the software owner's reproduction right.¹⁷⁸ Instead, the individual command

175. See 17 U.S.C. § 106(1) (2012).

176. See *id.* 106(2)–(6).

177. See *supra* Subsection II.B.2.

178. See *supra* Subsection II.B.2.

lines reproduced in the CPU, or cache registers close to the CPU, tend to be too fleeting to count as copies, too small to show originality, and too functional to constitute protectable expression.¹⁷⁹ Thus, in the context of cloud offerings, two copies can serve far more users than in the desktop context.

We must further consider whether users' viewing of GUIs on remote computers increase the copy count. In the cloud context, users can never see the software itself, which stays hidden on the cloud provider's server (on hard disk, in RAM, cache, and CPU). Many computer programs do not embody any meaningful GUI, so their analysis ends without any additional copies added to the count. For programs that do embody a meaning GUI, however, multiple remote users can see the GUI of the software. In response to access requests, the cloud provider sends copies of HTML code to display the software's GUI on each remote computer. Thus, there can be one GUI copy per remote user. Are more actionable reproductions thereby implicated? The answer is often negative, as such GUI copies may not "count" for purposes of copyright infringement. The first reason for this conclusion is that many GUIs are bare of copyrightable material. After all, GUIs tend to be highly functional and uniformly aligned with formats in the public domain to accommodate user expectations. But even when dealing with the exceptional GUI that reflects a creative contribution of expression, a cloud provider could avoid displaying it remotely by suppressing it. Specifically, the cloud provider need only prepare its own add-on GUI to mask the original GUI. This superimposition need not impair the software's functionality, as the GUI is not software per se, but rather is the output of software.¹⁸⁰

Thus, cloud scenarios permit a far greater number of users to utilize one software copy without requiring more copying more than in the traditional desktop scenario. Consequently, cloud scenarios do not implicate copyright owners' reproduction rights any more than in the desktop scenario, even when the user of a single software copy installs it on a cloud server for use by thousands of other users.

Cloud scenarios likewise often will not implicate the copyright owner's other rights.¹⁸¹ Regarding the adaptation right,¹⁸² cloud providers at times may have to modify code that was not originally written for cloud

179. *See supra* Subsection II.B.2.

180. *See supra* Subsection II.A.3–4.

181. For a more expansive treatment of these issues, see generally Lothar Determann, *What Happens in the Cloud: Software as a Service and Copyrights*, 29 BERKELEY TECH. L.J. 1095 (2014).

182. *See* 17 U.S.C. § 106(2) (2012).

deployment. However, if the provider can deploy the remote-access functionality with independently created programs, neither the cloud provider nor its end users would seem to implicate the software owner's adaptation right.¹⁸³ Moreover, whatever combinations or modifications may occur in RAM or CPU cache do not seem to reach sufficient levels of creativity or fixation to amount to adaptation.¹⁸⁴ Thus, if the software supplier delivers software in "cloud-ready" form or if the cloud provider can achieve "cloud-readiness" with independently created or licensed programs, the underlying programs' adaptation right should not be implicated. Turning to the distribution right,¹⁸⁵ the cloud provider does not transfer copies of the implicated software to the customer's computer—neither complete copies stored on ROM, nor partial copies stored in RAM or in cache. Instead, all copies remain on the cloud provider's server.¹⁸⁶ Accordingly, the cloud model should not implicate the distribution right of the underlying programs, either.

Regarding public performance rights,¹⁸⁷ the Act teaches that

[t]o "perform" a work means to recite, render, play, dance, or act it, either directly or by means of any device or process or, in the case

183. Specific cases might hinge on how much the underlying program had to be modified, as opposed to how much new expression the Cloud provider needed to achieve remote access functionality.

184. Arguably to the contrary of that perspective is *Dun & Bradstreet Software Services, Inc. v. Grace Consulting, Inc.* 307 F.3d 197, 204 (3d Cir. 2002) (holding that a software company distributing independently developed software modules for use with another software company's software program infringes the other company's copyrights because copies of the module and program are merged in computer memory and thus create a derivative work). But, this case does not seem to reflect a widely shared perspective. See generally Determann, *supra* note 1, at 1443 *et seq.* Moreover, if the case does state good law, its holding would equally affect software combinations in desktop-type deployments and therefore, the questions it raises are not peculiar to the Cloud.

185. See 17 U.S.C. § 106(3) (2012).

186. In the cloud context, customers do not even possess new copies of the code. Instead, those customers can only download the output that they create with the software (e.g., Word documents, PowerPoint slides, Excel spreadsheets) to their own computers. They cannot obtain copies of the software that runs on the cloud provider's servers to create the output (e.g., Microsoft Word, PowerPoint, or Excel application software). See Michael P. Widmer, *Application Service Providing, Copyright, and Licensing*, 25 J. MARSHALL J. COMPUTER. & INFO.L. 79, 95 (2007).

187. See 17 U.S.C. § 106(4) (2012). Note that this right is inapplicable to certain types of copyrightable works. But it does apply to "literary . . . works," and therefore encompasses software. The Copyright Act also embodies another right of public performance, applicable only to sound recordings. See *id.* § 106(6). As that subject matter is distinct from the instant subject, this Article confronts that matter no further. See 2 NIMMER & NIMMER, *supra* note 62, § 8.22.

of a motion picture or other audiovisual work, to show its images in any sequence or to make the sounds accompanying it audible.¹⁸⁸

The enumerated activities (recite, render, play, dance, act) all require that the work in question be presented to a human audience in a manner that can be visually or audibly perceived. By contrast, a computer's internal execution of code does not cause or allow perception by a human audience.¹⁸⁹ Therefore, a quick answer¹⁹⁰ is that the performance right of the underlying programs likewise should not be implicated by the cloud model. Finally, regarding the public display right,¹⁹¹ any material displayed remotely¹⁹² on users' screens (outside of the GUI, which can simply be replaced as discussed above)¹⁹³ is unlikely to amount to copyrightable expression sufficient to warrant an infringement cause of action.¹⁹⁴ Therefore, it seems overwhelmingly likely

188. 17 U.S.C. § 101 (2012).

189. Of course, software designed to run a video game could be performed. Likewise, software that embodies video or audio tutorials could see those aspects streamed, which in turn would create a public performance. The discussion in the text above is extremely cursory. Consider the following counterpoint:

At best, defendant Yeo's alleged publication of the ChainRxn video game for play by Facebook users constituted a public performance of plaintiff's copyrighted work under 17 U.S.C. § 106(4). Just as Congress considered the "reading a literary work aloud" as a performance rather than display of a literary work, the reading of Boomshine's copyrighted source or machine code by a computer (resulting in the presentation of the video game to the user) could be seen as an analogous performance of the underlying work. *See* H.R. REP. NO. 94-1476, at 63 (1976). Admittedly, this area of the law is still developing.

Miller v. Facebook, Inc., No. C 10-00264 WHA, 2010 WL 2198204 (N.D. Cal. May 28, 2010).

190. Bearing in mind the caveats of the previous footnote, the Ninth Circuit has held that merely playing interactive video games in public cannot implicate public performance rights because the concept of playing as performance had been narrowly interpreted to apply only to films and music. Allowing copyright owners games to control if and where games are played would unreasonably strengthen the copyright owners' interests at the expense of public interests in access to games. *See* Allen v. Academic Games League of Am., Inc., 89 F.3d 614 (9th Cir. 1996).

191. *See* 17 U.S.C. § 106(5) (2012).

192. Regarding the public display that occurs "at the place where the copy is located," such public display is immune from copyright liability under an extension of the first sale doctrine. *See id.* § 109(c).

193. *See supra* text preceding note 183.

194. U.S. courts have denied copyright protection for command line arrangements in office software products because these arrangements constitute methods of operation, which are excluded from copyright protection. *See* Lotus Dev. Corp. v. Borland Int'l, Inc., 49 F.3d 807 (1st Cir. 1995), *aff'd by an equally divided Court*, 516 U.S. 233 (1996). Similarly, commonly used icons and symbols lack sufficient originality or are dictated by extrinsic factors (user

that the cloud model does not implicate the underlying program's public display rights.

The foregoing considerations introduce new dynamics. In theory, they create a tremendous risk to software copyright owners, which may no longer have the power to force each user to acquire her own copy. Instead, the cloud provider may be able to simultaneously service many users. In the United States, at least, *Vernor* and its progeny may mitigate this risk by magnifying such owners' rights. To the extent that software purveyors continue to succeed in establishing that they only "license" those copies, they can simply insert a clause prohibiting cloud exploitation into the "license" terms. In that manner, contract law will step in to minimize their loss of rights under copyright law.¹⁹⁵

Nonetheless, these cloud-prohibition clauses do not shrink the concerns of software owners into insignificance. First, as we have seen above, European law regularly views software copies as being "sold" rather than "licensed." Therefore, the lenity of U.S. law cannot apply there.¹⁹⁶ Second, U.K. law magnifies users' freedom by not requiring copyright owners' consent to reproduce even highly creative websites or GUIs.¹⁹⁷ Third, danger to copyright owners is not altogether absent in the United States because not all courts follow *Vernor*. Moreover, the Supreme Court may reconsider *Vernor's* holding if the occasion arises.¹⁹⁸ Fourth, even if no chinks ever develop in *Vernor's* armor, a proprietor may surrender possession of a given copy, for whatever reason, outside of license terms.¹⁹⁹ A cloud purveyor would only need one copy to take advantage of the various features

expectations and familiarity). *See* *Apple Computer, Inc. v. Microsoft Corp.*, 35 F.3d 1435 (9th Cir. 1994).

195. Of course, to be efficacious, the subject terms must constitute a *condition* to the license, rather than a *covenant*. That distinction proved fatal to the copyright claim in *MDY v. Blizzard*, the companion case to *Vernor* that was referenced above. *See supra* note 141. Specifically, copyright owner Blizzard in that case sought to impose liability on counterdefendant MDY. It failed in that endeavor when the Ninth Circuit interpreted the provision under investigation to be a mere covenant, whose violation accordingly did not give rise to an infringement claim. *See* *MDY Indus., LLC v. Blizzard Entm't, Inc.*, 629 F.3d 928, 939–42 (9th Cir. 2010).

196. *See supra* Subsection II.C.2.

197. *Id.*

198. *See supra* Subsection II.C.1.

199. One could imagine a beta copy that was not shrinkwrapped, settlement of litigation that leaves a copy in a remote party's hands, a mistake at the factory that causes product to be shipped absent its shrinkwrap terms or printed defectively, or a variety of other factors.

catalogued above without triggering copyright infringement liability, despite making the third-party's software accessible to thousands of third parties.²⁰⁰

IV. THE CLOUD'S IMPACT ON SOFTWARE COPYRIGHTS

Based on the above review of past challenges relating to software and copyrights, along with the development of cloud exploitation models, the stage is set to assess the cloud's impact.

A. CHALLENGES TO COPYRIGHT LAW'S GOALS

We have seen above how an aggressive implementation of *Vernor* could set at naught the delicate balance underlying software protection. In particular, it could undermine the first sale doctrine, the essential step defense, and the application of fair use to software.²⁰¹ Still, the discussion has also noted that aggressive implementation might not pass legal muster—in other words, the Supreme Court might halt *Vernor's* advance before matters reached such a pass, at least in the context of traditional book publishing.²⁰²

Nonetheless, one must frankly concede that no such ameliorative construction seems likely in the cloud context. Each possible defense is examined in turn.

1. *First sale doctrine*

A software copyright owner can avoid the first sale doctrine by commercializing its software as a service rather than as a sale. Cloud service transactions involve recurring payments, temporal-use limitations, and no transfer of possession. Such transactions do not resemble sales by any standard. Besides moving to services models, software copyright owners have a few other options to reduce the risk that the first sale doctrine applies, even under E.U. laws. Software companies can include “sales-unlike” clauses in contracts,²⁰³ apply technological restrictions,²⁰⁴ charge extra for re-sales or

200. Other pitfalls also loom, such as the distinction between covenant and condition if the operative cause of action is for copyright infringement, and lack of privity if the operative cause of action is breach of contract. *See supra* note 195.

201. *See supra* Section II.B.

202. *See supra* Subsection II.C.3.a.

203. *See* Determann & Batchelor, *supra* note 118 (noting that software companies can consistently use terminology to clarify that they are selling licenses, services, or access to software, rather than software copies. They can include “unlike-sales” commercial terms, “field-of-use” restrictions, a contractual obligation to return old software copies at the time of upgrades, access limitations (authorized or concurrent users), and others).

204. Software that is frequently updated, upgraded, and changed, without reverse version interoperability, is more difficult to resell—but may also be less attractive to users. Dongles and expiring activation codes can also be used to control changes on the end-user

transferable licenses, charge more for support on transferred licenses, grant “real” reproduction licenses that cannot be transferred based on the exhaustion principle (e.g., enterprise licensing), etc. However, with a shift to cloud models, copyright owners could preclude secondary markets more clearly and effectively.

This state of affairs generally corresponds to *Vernor*'s ruling. Nonetheless, the differences are as important as the points of commonality. First, and most importantly, Europe applies the opposite construction.²⁰⁵ Nonetheless, the cloud would Americanize the content of European law in this regard. By moving to cloud business models, a software copyright owner could continue to control distribution of its software and side-step any first sale under E.U. law.

Second, internal to United States law, the cloud eliminates any conceptual cloud hanging over *Vernor*'s construction. As previously explicated, the problem with the Ninth Circuit's logic inheres in the Ninth Circuit's construction that a proprietor's permanent parting with physical ownership for a one-time fee does not constitute a sale.²⁰⁶ Nonetheless, even *Vernor*'s harshest critics concede that the cloud does not result in a sale of physical products.²⁰⁷ In other words, when a user does not obtain permanent dominion over CD-ROMs containing AutoCAD, but instead rents access to that software product by the hour, all parties agree that no “sale” has taken place. Accordingly, the cloud itself affords no room to vindicate the ruling of earlier cases to the contrary of *Vernor*.²⁰⁸

Third, the aggressive extension of *Vernor* under current copyright law to books, as postulated above, seemed dicey at best.²⁰⁹ By contrast, purveying ebooks over the cloud is straightforward and beyond challenge. In the ebook context, users do not gain dominion over physical products that the first sale doctrine would allow them to resell.²¹⁰ Accordingly, a user privilege that traces back to 1908 and is currently codified in § 109 of the Act, could effectively be eliminated. The upheaval to traditional notions of copyright law could not be more pronounced.

side, regarding hosting equipment, authorized users, and other details—but such restrictions are not favored by end users or data privacy laws.

205. See *supra* Subsection II.C.2.

206. See *supra* Subsection II.C.1.

207. See 2 NIMMER & NIMMER, *supra* note 62, § 8.12[B][1][d][iii].

208. See *supra* note 97.

209. See *supra* Section II.C.

210. See 2 NIMMER & NIMMER, *supra* note 62, § 8.12[E] (no digital first sale defense).

On the other hand, a move towards embracing the first sale doctrine as radically as *UsedSoft* suggests could cause the cloud to extremely marginalize the software copyright owner's rights and economic opportunities. Under *UsedSoft*, once a software developer parts with one software copy under an agreement characterizing the transaction as a non-exclusive, limited license, a licensee-turned-owner could deploy the copy over the cloud and offer it to thousands of other users in competition with the software developer.²¹¹ If U.S. courts do not follow *UsedSoft* or employ European views on copy ownership in cross-border cases,²¹² cloud providers might find a home in Europe for cloud offerings. These providers might deliver cloud offerings remotely to users in the United States without actually causing copies to be made in the United States, disrupting legitimate U.S. copyright exploitation in the process.

2. *Essential step defense*

The previous subsection illustrated how copyright owners can unilaterally eliminate users' distribution privileges by purveying their wares on the cloud. Different considerations apply if this state of affairs is extended to users' reproduction privileges in software. The aggressive application of *Vernor* postulated above imagined that copyright owners of software would charge \$490 to "license" a physical copy of software, and then impose an additional charge to use it (e.g., by loading it into RAM). How do those aspects translate to the cloud?

As a theoretical matter, danger exists that this aspect of users' privileges will equally be forfeited. As a practical matter, however, there is reason to be more sanguine. Granted, cloud purveyors can impose any charges that the market will bear—for instance, \$2 for the first four hours of using Program X, and then \$1000 for every minute thereafter. Yet it is doubtful that purveyors will be able to charge for "licensing" a physical copy of the software, only to impose a hidden "use" charge later. After all, there is no physical copy provided at the outset to entice users to pay. In other words, the charge will most likely be for use, not licensing. The same potential for abuse ventilated above seems to be largely absent regarding the essential step defense.

3. *Reverse engineering*

An aggressive application of *Vernor* also elicits different considerations when applied to the fair use defense's protection of reverse engineering. In

211. See *supra* Section III.C.

212. See *Importing Software and Copyright Law*, *supra* note 114, at 4–5.

this context, reverse engineering means the process of copying, analyzing, and disassembling object code in order to separate unprotectable, functional elements from a copyrightable expression. The cloud, however, appears to make traditional forms of reverse engineering impracticable.²¹³ This state of affairs results not so much from limitations established by copyright law, but rather from the cloud provider's technological ability to control and restrict access to the underlying code residing on the cloud provider's server.²¹⁴ In other words, in the cat-and-mouse competition between software publishers and competitors, the latter have previously been able to reverse engineer, given the business models and technical controls available. The advent of the cloud promises to overturn this power imbalance.

Of course, future competitors could attempt self-help by removing the software publishers' technical measures and attempting to reverse engineer the product anyway. Unfortunately for them, however, those who attempt to reverse engineer software underlying cloud offerings face civil and criminal penalties under laws prohibiting unauthorized access to computers and circumvention of technical protection measures. These protections and penalties are codified in the Digital Millennium Copyright Act,²¹⁵ the Computer Fraud and Abuse Act,²¹⁶ and equivalent laws in the E.U.²¹⁷

These same kinds of laws—protecting computers against interference—can be used to both prevent the creation of interoperable software products through reverse engineering, as well as establish absolute hurdles against interoperability of independently created software programs. In the age of the cloud, software offered as a service on one computer often may access software offered as a service on another computer in order to communicate data. Such access can be hindered based on laws restricting computer interference and circumvention of technical protection measures.²¹⁸ Courts have rejected arguments by defendants that the fair use doctrine under copyright law must allow them access to hosted software copies, in

213. See Craig Zieminski, *Game Over for Reverse Engineering?: How the DMCA and Contracts Have Affected Innovation*, 13 J. TECH. L. & POL'Y 289, 307–09 (2008).

214. In addition, the pertinent contracts promulgated by the cloud provider can be to the same effect.

215. See 17 U.S.C. § 1201 (2012).

216. See 18 U.S.C. § 1030 (2012).

217. See Ian Walden, *COMPUTER CRIMES AND DIGITAL INVESTIGATIONS* (2007); see also Lothar Determann, *Internet Freedom and Computer Abuse*, 35 HASTINGS COMM. & ENT. L.J. 429 (2013) (discussing the history, context and current reform proposals regarding computer interference laws in the United States and abroad).

218. See *MDY Indus., LLC v. Blizzard Entm't, Inc.*, 629 F.3d 928, 945 (9th Cir. 2010); *Craigslist Inc. v. 3Taps Inc.*, 942 F. Supp. 2d 962 (N.D. Cal. 2013) (finding violations of the DMCA's anti-circumvention measures by automated access to websites).

derogation of the website operator's terms of use or in violation of technical protection measures, even if the online access was for the purpose of developing or deploying interoperable applications.²¹⁹ Therefore, cloud providers can largely rely on technical protection measures and contractual website access restrictions to protect the underlying code. Once such measures and access restrictions are in place, others cannot access the software underlying the cloud offering, except in compliance with the authorizations contained in the applicable contracts.

Therefore, the move to cloud models may potentially result in less interoperability and adverse effects on innovation. Whether this potential will materialize remains to be seen. Currently, many cloud platforms encourage the development of compatible applications and have spurred a flurry of development.²²⁰ On the other hand, frictions have developed and software copyright owners with established platforms have been able to prevent the development of add-on offerings that they probably could not have achieved under more traditional distribution models.²²¹ If this situation becomes a problem and significantly stifles innovation, changes to the legislative framework may have to be considered, including similar defenses for access and reverse engineering of software for interoperability purposes in the cloud, as in more traditional distribution models.

B. OTHER IMPACTS

Of course, the cloud also poses challenges to other aspects of copyright law.

219. See *MDY*, 629 F.3d at 950; *Ticketmaster L.L.C. v. RMG Techs., Inc.*, 507 F. Supp. 2d 1096 (C.D. Cal. 2007) (development, hosting, distribution, and/or use of software designed to access Ticketmaster website for bulk ticket purchases in circumvention of contractual and technical access restrictions constitutes copyright infringement in Ticketmaster's web pages and violation of the Computer Fraud and Abuse Act); *Facebook, Inc. v. Power Ventures, Inc.*, 844 F. Supp. 2d 1025, 1028 (N.D. Cal. 2012) (automated access to Facebook pages for marketing purposes in violation of Facebook terms of use constitutes a violation of the Computer Fraud and Abuse Act and copyright infringement); Lothar Determann & Irene Gutierrez, *Copyright Violations in Caching Website Content, Contract Formation*, 3 J. INTEL. PROP. L. & PRAC. 548 (2008); Ziemiński, *supra* note 213.

220. See, e.g., *APPLE APP STORE*, <http://www.apple.com/iphone/from-the-app-store/> (last visited December 13, 2014).

221. See, e.g., *Ticketmaster*, 507 F. Supp. 2d 1096; Determann & Gutierrez, *supra* note 219; *MDY*, 629 F.3d at 945.

1. *Open-source-license terms*

Today, most commonly used open-source license terms permit reproduction and adaptation freely and unconditionally.²²² In order to preserve developers' freedom to tinker with code, the license terms typically refrain from applying any restrictions until the developer distributes a copy to third parties.²²³ Consequently, companies that operate service businesses have been able to use most open-source code without being legally obligated under the applicable license terms to give back to the community.²²⁴ Cloud models sidestep some of the attempts by the open-source movement to keep software free and available. In cloud models, however, the code is locked up on servers and not available for further improvement and development.

This matter presents a choice to the drafters of open-source license terms. Licensors can tie release obligations or other restrictions not only to distribution, but also to offering modified or unmodified software as a service, given that cloud offerings always implicate reproduction rights.²²⁵ The GNU GPLv3 (Affero version) already embodies such a provision²²⁶ and other licenses may follow suit.²²⁷ Thus, the move to service models may prompt some open-source code licensors to consider updating their license terms. In the meantime, companies can sidestep most commercial concerns relating to open or free software code by switching from traditional forms of software distribution to cloud models.

2. *Unauthorized Access to Copyrighted Material*

a) Piracy

Software pirates around the world blatantly copy literal code without the copyright owner's authorization.²²⁸ This problem significantly afflicts software commercialization models that are based on delivery of software

222. See *The Open Source Definition*, OPEN SOURCE INITIATIVE, <http://opensource.org/osd/> (last visited Dec. 13, 2014) ("Free Redistribution: The license shall not restrict any party from selling or giving away the software The license must allow modifications and derived works")

223. See, e.g., *Sections 0 and 2*, GPL, <http://www.gnu.org/copyleft/gpl.html> (last visited Dec. 13, 2014).

224. See Theresa Gue, *Triggering Infection: Distribution and Derivative Works under the GNU General Public License*, 2012 U. ILL. J.L. TECH. & POL'Y 95 (2012).

225. See *supra* Part III.B.

226. *GNU Affero General Public License*, GNU OPERATING SYS., <http://www.gnu.org/licenses/agpl-3.0.html> (last visited Dec. 13, 2014).

227. Gue, *supra* note 224, at 101.

228. BUS. SOFTWARE ALLIANCE, *SHADOW MARKET: 2011 BSA GLOBAL SOFTWARE PIRACY STUDY* (9th ed. 2012), available at <http://portal.bsa.org/globalpiracy2011/>.

copies on physical media or by way of download.²²⁹ In the cloud context, pirates find it much more difficult to obtain software copies, given that the code remains on heavily secured servers and is not widely disseminated.²³⁰ In the best case scenario, cloud offerings will undermine much of the very incentive underlying piracy, as users are afforded the ability to obtain the works they want when they want from wherever they are located—all for (hopefully) a reasonable price.²³¹

Nonetheless, the software copyright owners' economic interests can be adversely affected by cloud customers' sharing of access credentials in violation of cloud agreements, or by criminals hacking into cloud systems.²³² At present, those activities seem easier to prevent and prosecute than literal reproduction of physical software copies. Therefore, hosted software in cloud models tends to be far less vulnerable to software piracy than in more traditional distribution models.²³³

b) Facilitation of Multiple Cloud Users

Copyright owners are separately threatened by the cloud's ability to disseminate software to numerous third parties. A user who obtains one physical copy of that software can offer access to it, via the cloud, to tens of thousands of third parties, thus obviating the need for any of those myriad end-users either to obtain their own copy or to access the proprietor's own cloud offering. Viewed through that prism, the cloud portends disaster to those copyright owners. Any positive influence the cloud has on rightsholders' bottom line through universalizing the effects of *Vernor* are more than counterbalanced by the dangers of cloud exploitation escaping copyright liability.²³⁴

229. See, e.g., *Piracy*, MICROSOFT, <http://www.microsoft.com/en-us/piracy/default.aspx> (last visited December 13, 2014).

230. See Matt Asay, *The BSA's Fading Twentieth-Century Piracy Fight: Misreading the Data*, REGISTER (Sept. 24, 2010), http://www.theregister.co.uk/2010/09/24/piracy_open_source_bsa/.

231. "Combining cloud computing and content streaming technologies could thus reduce online piracy of entertainment content by providing the consumer with value—the ability to access content from almost anywhere—while providing content owners, creators, and providers with an unprecedented means to control their digital works." Tamara Celine Winegust, *Work with Your Head in the Clouds: The Impact of Cloud Computing and Content Streaming on Copyright in the Entertainment Industry*, 4 AM. U. INTELL. PROP. BRIEF 8, 10 (2012).

232. See Florence de Borja, *Can Cloud Computing Stop Software Piracy?*, CLOUD TIMES (Sept. 7, 2012), <http://cloudtimes.org/2012/09/07/can-Cloud-stop-software-piracy/>.

233. See Asay, *supra* note 230.

234. See *supra* Section III.C.

As noted above, the legal underpinnings of that model have yet to be fully clarified.²³⁵ Much may depend on the precise technical specifications employed by various cloud purveyors, as well as how many distinct copies they need to generate to service their customers. At present, all that can be stated is that dangers and opportunities abound for all interested parties.

V. SOME WEATHER FORECASTS

With the benefit of the foregoing exposition, it is time to evaluate the cloud and its implications for copyright on software and beyond. Given that the phenomenon is still at its outset, the time is premature for a marriage counselor to conclude either that there are irreconcilable differences or that a bit of effort on both parties' sides will ensure another three decades of at least tolerable cohabitation. The savvy observer must simply register current tensions in light of past history, extending strictly tentative diagnoses regarding the balms needed for future non-violent relations.

A. GLOOMY PROSPECTS: DARK CLOUDS UNDERMINE THE GOALS OF 1980

RESOLVED, The cloud is unprecedented, both in its technology and the stress that it places on traditional copyright categories.

Congress embraced software within the copyright domain, *faute de mieux*, and subject only to a carefully developed balance: at the same time that it conferred rights on copyright under § 102 of the Copyright Act, it adopted the limitations on the distribution right in § 109; equally, Congress added new limitations on the reproduction right under § 117, acting against the backdrop of copyright law's ubiquitous defense, the infinitely malleable fair use doctrine.

As set forth above, certain interpretations of existing law (celebrated by some, bemoaned by others) yield a tally of Copyright Owners 4, Users 0.²³⁶ For those who bemoan those results, at least current law contains some prospect for amelioration. Given that these critics decry *Vernor's* construction of the statute as faulty, they hold onto the hope that other circuits, eventually joined by the Supreme Court, will ultimately jettison that ruling and set the law back on its proper course.

No such hope arises, however, as to the cloud. Even *Vernor's* critics concede that those who purvey software via the cloud may invoke existing law to eliminate any user privileges of further distribution. Moreover, they

235. *See supra* Section III.C.

236. *See supra* Subsection II.C.4.

may equally deny any user privileges of further reproduction, albeit that particular danger is noticeably muted in the cloud context. Finally, cloud offerings effectively lie beyond the realm of fair use exploitation. As such, all use becomes “fared,”²³⁷ with no latitude remaining for one of the fundamental protections encapsulated in the text of the 1976 Act itself.²³⁸

These deformations push the law so far out of its intended path that reformation becomes essential. As the Vice-Chair of CONTU commented in the report on which Congress relied when extending copyright protection to software in 1980, certain “line[s] of demarcation” must be borne in mind, not because they were needed as of 1978 when the CONTU report issued, but rather because they “may prove useful in the years to come if the current recommendation for protection of all software should prove unduly restrictive.”²³⁹

That time has now dawned. The cloud is its midwife.

In fact, the cloud’s deformation is much worse than even the Vice-Chair of CONTU imagined three-and-a-half decades ago. For not only does the cloud eliminate the first sale doctrine as to software, but it also exerts the same effect across the field of all literary works. Since 1908, *Bobbs-Merrill Co. v. Straus* and its statutory codification have safeguarded the privilege for readers of novels to resell them at second-hand bookstores. In a world of ebooks, that privilege is no more.

The same deprivation applies across the board, moreover, beyond the sphere of literary works. Video rental stores have traditionally offered movies to their customers. That instantiation serves for audiovisual works the role that libraries have traditionally served for literary works. Those outlets will be similarly barred in the cloud’s universe. And the same applies to music and other works subject to copyright protection as well. In fact, the only secondary market for works of authorship, in a future dominated by the cloud, promises to be in the realm of fine art. Thus, purchasers of a painting to hang on their living room wall will be able to resell the work, as will the sculpture garden that wishes to cycle out an old maquette for a new bronze.²⁴⁰ Outside those peripheral applications, however, the first sale doctrine will be rendered otiose everywhere.

237. See Tom W. Bell, *Fair Use vs. Fared Use: The Impact of Automated Rights Management on Copyright’s Fair Use Doctrine*, 76 N.C. L. REV. 557 (1998).

238. See *supra* note 165.

239. See FINAL CONTU REPORT, *supra* note 29.

240. But even those resales will become more complicated, to the extent that the Copyright Act is amended to adopt the *droit de suite*, a prospect currently under consideration. See Resale Royalty Right, 78 Fed. Reg. 19326 (Mar. 25, 2013).

Thus the dream of those “license” advocates²⁴¹ who, long ago, rendered the first sale doctrine a dead letter will be universalized. In their wake, Fred Hoyle’s dystopian vision becomes our reality.²⁴² To save us from that darkness, legislative redress is needed.

B. A CHEERY REBUTTAL: RETURN TO AN EARLIER SENSIBILITY

RESOLVED, The cloud is not unprecedented and simply returns many traditional copyright categories to their historic role.

On the other hand, a broader view of history yields different insights. We change our field of vision here from the “literary works” that have occupied this Article thus far—whether those denominated *software* or more traditional forms, such as *novels*—in order to scrutinize the realm of music. Here, we reach different territory.

Whereas *Bobbs-Merrill Co. v. Straus* litigated the status of *The Castaway* over a century ago, at a time when any interested reader could obtain her own personal copy of the work in question, matters were simultaneously very different for music lovers.

Throughout most of history, music has been experienced exclusively as a service, or something that one needs to receive from a service provider in order to enjoy. The advent of sound recordings—especially digital sound recordings—commoditized the medium, allowing consumers to take music with them wherever they go, as well as giving them the ability to manipulate and engage with the sound recordings at will. “Internet music streams, however, remove consumers’ control over the access and playback of music, transforming digital music once again into a service.”²⁴³

The same considerations apply to movies. The motion picture industry was in its infancy when *The Castaway* hit bookstores. For many decades thereafter, the only way to experience a movie was to wait until it came to one’s local cinema, or later when a local television station carried its broadcast.²⁴⁴ The situation changed only with the advent of the Betamax in the 1970s.

We can therefore appreciate that, before high-fidelity music recording and videotape recording capabilities, users had none of the privileges

241. See *supra* text accompanying note 101.

242. See generally FRED HOYLE, *THE BLACK CLOUD* (William Heinemann Ltd. ed., 1957) (positing intelligence circling earth in a cloud that blocks out the sun).

243. Jay Anderson, *Stream Capture: Returning Control of Digital Music to the Users*, 25 HARV. J.L. & TECH. 159, 162 (2011).

244. See David Nimmer, *Brains and Other Paraphernalia of the Digital Age*, 10 HARV. J.L. & TECH. 1, 14–16 (1996).

canonized above. During those intervals, they had no ability to obtain ownership of songs or movies; instead, they had to wait for others to perform them. In addition, they certainly enjoyed no “essential step” defense in that context. And their rights of fair use were also extremely circumscribed, given no practical way of quoting at length from the music and movies that had been evanescently performed and then passed into the ether.²⁴⁵

Viewed from this perspective, the cloud simply returns music to its origins,²⁴⁶ and the same for film too. By replacing products with streams, the cloud brings back to the fore the effective copyright status of large swaths of works that existed for many decades in the past.

If the enjoyment of movies in the 2030s matches that of the 1930s—in which the proprietor could make a charge for each and every viewing²⁴⁷—no cosmic injustice requires redress. Instead, the long eye of history can take bemusement in the swirls and eddies of time, in which nothing is really new under the sun.

C. HIGHS AND LOWS IN TURBULENT TRANSITIONS

Whatever the long-term weather forecast may hold for clouds, there is one prediction which we do not hesitate to make at this juncture: there will be turbulent transitions.²⁴⁸ Software in the cloud places entirely different pressure points on copyrights than software on desktops. Phases of high pressure will alternate with low-pressure systems moving in as the marketplace, technological measures, courts, and perhaps even legislatures adjust. But any future envisioned today will inevitably be upset by tomorrow’s reality. We have no choice but to live and learn.

245. As to movies, there was no practical way to “perform” them at home. As to music, by contrast, anyone with the requisite skill was welcome to play the tune on the living room piano. Presumably, even a neighborhood gathering to sing familiar tunes would enjoy a robust fair use defense.

246. “Music began as a service and has remained exclusively so for most of its existence. Until very recently in human history, a person wishing to hear music required the performance of a musician.” Anderson, *supra* note 243, at 162.

247. In the twentieth century, the charge from the studios was either to theaters or to broadcast networks. In the current century, the charge will be directly to consumers—to whom the charge was previously passed on by those intermediaries.

248. High honors go here to the prospects of cloud purveyors offering proprietary software without authorization to thousands of remote users, as discussed above. *See supra* Section III.C.

SOVEREIGNTY UNDER SIEGE: CORPORATE CHALLENGES TO DOMESTIC INTELLECTUAL PROPERTY DECISIONS

Cynthia M. Ho[†]

ABSTRACT

Countries face a new threat that strikes at their ability to balance protection of intellectual property rights against other priorities, such as public health. They may have to pay substantial compensation to companies that dislike domestic intellectual property laws. This threat is much more significant than the Agreement on Trade-Related Aspects of Intellectual Property Rights (“TRIPS”), a landmark international agreement concluded twenty years ago, that for the first time required all countries to provide “minimum” levels of intellectual property rights; before that time, countries were not obligated to provide any such rights at all. Since the conclusion of TRIPS, policymakers and scholars have strived to preserve local flexibilities to consider domestic policies, such as public health. However, those flexibilities may quickly evaporate if companies can bring claims against countries for compromising private investments under so-called “investor-state arbitration” claims.

This is not a theoretical problem—Eli Lilly is currently seeking \$500 million in compensation from Canada because Canadian courts invalidated two of its patents under prevailing law. In addition, there are unique issues raised by Eli Lilly’s claim that transcend broader concerns raised by scholars and commentators concerning investor-state disputes. In particular, if Eli Lilly’s claim succeeds, it will disrupt internationally accepted norms that permit countries to have different standards of protection. This Article provides a detailed analysis of Eli Lilly’s case of first impression. In so doing, the Article both explains why an arbitration tribunal should reject Eli Lilly’s claims, and predicts the likely impending threats to domestic regulation of public health that intersect with the interests of pharmaceutical companies. This Article ultimately proposes specific language to incorporate in pending agreements to forestall such predicted harms.

© 2015 Cynthia M. Ho

† Clifford E. Vickrey Research Professor, Loyola University Chicago School of Law. I would like to thank Karen Cross, Christophe Douaire de Bondy, Susan Franck, James Gathii, Christi Guerinni, Yaniv Heled, Molly Beutz Land, Margaret Moses, and Jerome Reichman for comments or other assistance on earlier drafts, as well as participants at IPSC 2014. In addition, many thanks to the assistance of librarian Patricia Scott, as well as research assistant Jean Liu for their excellent help.

TABLE OF CONTENTS

I.	INTRODUCTION	215
I.	BACKGROUND	225
A.	DOMESTIC PATENT LAW AND POLICY	225
B.	THE INTERNATIONAL LANDSCAPE: IP AND FOREIGN INVESTMENTS	227
1.	<i>Intellectual Property: Territoriality and Treaties</i>	228
2.	<i>International Agreements Governing Foreign Direct Investments and Providing Investor-State Disputes</i>	231
C.	ELI LILLY'S CASE	236
II.	REVOCATION OF INTELLECTUAL PROPERTY RIGHTS SHOULD NOT BE A BREACH OF INVESTMENT OBLIGATIONS	242
A.	CANCELED INTELLECTUAL PROPERTY RIGHTS SHOULD NOT BE A COVERED INVESTMENT THAT WOULD ENABLE AN INVESTOR-STATE DISPUTE	243
1.	<i>Intellectual Property is Different from Real Property</i>	243
2.	<i>Canceled Intellectual Property Should Be Excluded from Investment Agreements Based on Policy Grounds</i>	246
a)	International Agreements Permit Nations to Decide on the Scope of Intellectual Property Rights; Investment Arbitrations Should Not Disrupt This Norm	247
b)	Domestic Policy Underlying Intellectual Property Supports Deferring to States	252
c)	Considering Intellectual Property an Investment Does Not Foster Traditional Policy Justifying Investor-State Disputes	255
B.	INVALIDATION OF PATENT RIGHTS SHOULD NOT CONSTITUTE EXPROPRIATION	258
1.	<i>Canada's Actions May Be Exempt from an Expropriation Claim</i>	259
2.	<i>Canada Could Be Found to Not Have Expropriated Eli Lilly's Patents</i>	263
a)	Canada Should Not Be Found to Have Directly Expropriated Eli Lilly's Patents	263
b)	Canada Should Not Be Found to Have Indirectly Expropriated Eli Lilly's Patents	264
C.	INVALIDATION OF PATENT RIGHTS SHOULD NOT CONSTITUTE A VIOLATION OF THE FAIR AND EQUITABLE TREATMENT STANDARD	272

1.	<i>Eli Lilly Has No Legitimate Expectation that Common Law Interpretations Will Not Change</i>	276
2.	<i>An Issued Patent Is Not a State Representation of Permanent Validity That Can Be Justifiably Relied On and Must Be Balanced against State Interests</i>	278
a)	A Patent Is Not a State Representation of Guaranteed Validity	279
b)	There Has Been No Negative Reliance upon a State Representation	281
c)	Eli Lilly Has No Legitimate Expectation That Outweighs Canada’s Interests	282
III.	BEYOND ELI LILLY’S CASE: PENDING PROBLEMS AND HOW TO ADDRESS THEM	283
A.	PUBLIC HEALTH ISSUES IN DANGER OF DISRUPTION.....	283
1.	<i>Patentability Standards and Compulsory Licenses</i>	283
2.	<i>Compulsory Licenses</i>	285
3.	<i>Domestic Regulation of Clinical Data</i>	289
a)	Countries That Do Not Provide “Data Exclusivity” Will Likely Be Challenged	289
b)	Domestic Data Transparency Requirements Are Vulnerable to Challenge	293
B.	PROPOSALS TO PRESERVE FLEXIBILITY UNDER TRIPS.....	297
1.	<i>Exclude Intellectual Property from Investor-State Disputes</i>	298
2.	<i>Limit the Scope of Investment Claims Based on International Agreements Such as TRIPS</i>	300
a)	Limit Expropriation Claims.....	300
b)	Limit Fair and Equitable Treatment Claims	302
IV.	CONCLUSION	303

I. INTRODUCTION

Is a company entitled to compensation from a country that declines to provide that company an intellectual property right? Eli Lilly thinks so. Eli Lilly is seeking \$500 million from Canada pursuant to an international agreement that permits foreign—but not domestic—investors to bring “investor-state arbitration” claims before a panel of private arbitrators against countries that interfere with its “investments.”¹ This is the first time a company has initiated an investor-state arbitration to challenge a domestic patent law that arguably complies with international agreements governing patents. In particular, Eli Lilly claims that two of its patents were

1. Alternatively, these are referred to as investor-state dispute settlements (“ISDS”).

investments, and that Canada unduly interfered with these investments when the country's courts invalidated the patents for failing to meet a Canadian patentability requirement.²

Eli Lilly's demand for substantial financial compensation discourages countries from fine-tuning their patent laws, even when domestic laws would comply with separate international agreements concerning intellectual property. International agreements that require protection of patents permit some domestic flexibility to promote social policies other than innovation, such as access to lower-cost medicine. However, when viewed in the broader historical context, the flexibility is in fact significantly limited and further underscores why investor-state disputes pose a particularly significant threat to sovereign rights to protect traditional domestic goals, such as public health. To understand the chilling effect, a brief background of the most important international agreement governing intellectual property and its significance follows. In 1994, over one hundred and twenty countries agreed for the first time to provide "minimum" standards of patent protection pursuant to the Agreement on Trade-Related Aspects of Intellectual Property Rights ("TRIPS").³ This was a major change for many countries. Before TRIPS, some countries were reluctant to grant patents on drugs because such patents would inevitably result in higher costs and less access to affordable medicine.⁴ For example, some only permitted inventors to patent methods of

2. *Eli Lilly & Co. v. Canada*, Notice of Arbitration (NAFTA/UNCITRAL Arb. Trib. Sept. 12, 2013) [hereinafter *Eli Lilly Notice of Arbitration*].

3. Agreement on Trade-Related Aspects of Intellectual Property Rights art. 27, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 299 [hereinafter TRIPS]; *The 128 Countries That Had Signed GATT by 1994*, WORLD TRADE ORGANIZATION, http://www.wto.org/english/thewto_e/gattmem_e.htm. However, WTO member countries that the UN designates Least Developed Countries do not need to provide patent protection on drugs until at least 2016. Decision of the Council for TRIPS of June 27 2002, *Extension of the Transition Period Under Article 77.1 of the Agreement for Least Developed Country Members for Certain Obligations with Respect to Pharmaceutical Products*, para. 1, IP/C/25 (July 1, 2002). However, there are pending proposals for the WTO to further extend this timeline. *E.g.*, Request for an Extension of the Transitional Period under Article 66.1 of the TRIPS Agreement for Least Developed Country Members With Respect to Pharmaceutical Products and for Waivers from the Obligation of Articles 70.8 and 70.9 of the TRIPS Agreement, *Communication from Bangladesh on behalf of the LDC Group*, IP/C/W/605 (Feb. 23, 2015).

4. *E.g.*, SUDIP CHAUDHURI, THE WTO AND INDIA'S PHARMACEUTICALS INDUSTRY: PATENT PROTECTION, TRIPS, AND DEVELOPING COUNTRIES 59 (2005). Indeed, Switzerland, home to major multinational pharmaceutical companies such as Novartis, only patented drug products in the late 1970s. PATENTVERORDNUNG [PatV], Patent Regulation [PatV] Oct. 19, 1977, SR 232.141 (Switz.). The Swiss Constitution originally barred any patents on drugs and only later were patents permitted on chemical processes. *See, e.g.*, MICHELE BOLDRIN & DAVID K. LEVINE, AGAINST INTELLECTUAL MONOPOLY 216 (2008).

making drugs, and not the drugs themselves, to ensure that drugs would not be subject to a patent premium.⁵ After centuries of limited patent protection of drugs, most countries must now provide patents on drugs pursuant to TRIPS. However, TRIPS notably only requires minimum, but not uniform, standards of protection, such that countries still have some flexibility to tailor patent standards to their respective interests.⁶ Since the conclusion of TRIPS, scholars and policymakers have recommended that developing countries with limited resources embrace TRIPS flexibilities to ensure that citizens in these countries are not unduly harmed by the inevitably higher cost of patented drugs.⁷ Although some developing countries recently considered modifying their laws to take advantage of these flexibilities, they may now have second thoughts—especially if the arbitration tribunal finds in Eli Lilly's favor.⁸

Eli Lilly's challenge impacts countries at a critical juncture. For more than a decade, the pharmaceutical industry has suffered from an innovation crisis in conjunction with the expiration of patents on highly profitable drugs, resulting in a struggle to sustain revenue.⁹ In response, the industry has increasingly patented drugs that are merely minor variations of existing drugs

5. After all, if only one company could make a drug, that company would likely charge high prices. On the other hand, if multiple companies can make the same drug, albeit with different methods, that should cause competition that reduces drug costs.

6. TRIPS, *supra* note 3, art. 1; *see also infra* Subsection II.B.1 (discussing flexibility under TRIPS).

7. *See infra* notes 150, 157.

8. *See infra* notes 149–50 and accompanying text (discussing Brazil and South Africa).

9. *E.g.*, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT ON PROPELLING INNOVATION IN DRUG DISCOVERY, DEVELOPMENT, AND EVALUATION 9–10 (2012); Fabio Pammolli et al., *The Productivity Crisis in Pharmaceutical R&D*, 10 NATURE REVIEWS DRUG DISCOVERY 428, 429 (2011); Charlotte Harrison, *The Patent Cliff Steepens*, 10 NATURE REVIEWS DRUG DISCOVERY 12, 12–13 (2011); Katie Thomas, *U.S. Drug Costs Dropped in 2012, but Rises Loom*, N.Y. TIMES, Mar. 18, 2013, at A1. Although some have contested that there is no innovation crisis, they are outliers. *See, e.g.*, Donald W. Light & Joel R. Lexchin, *Pharmaceutical Research and Development: What Do We Get for All That Money?*, BRIT. MED. J., Aug. 7, 2012, at 1 (claiming the innovation crisis is a myth used as a ploy to get government benefit, but nonetheless arguing that there is a crisis with respect to an inadequate number of new drugs that represent a true therapeutic advance); Bernard Munos, *Lessons from 60 Years of Pharmaceutical Innovation*, 8 NATURE REVIEWS DRUG DISCOVERY 959, 959–60 (2009) (concluding no decline in productivity, using an unduly restrictive data set). Indeed, many have noted that innovation has been stagnant for years. *E.g.*, Jeff Hewitt et al., *Beyond the Shadow of a Drought: The Need for a New Mindset in Pharma R&D*, OLIVER WYMAN 1, 3–4 (2011), http://www.oliverwyman.com/insights/publications/2011/nov/beyond-the-shadow-of-a-drought--the-need-for-a-new-mindset-in-ph.html#.VEP2e_nF90A; *Measuring the Return From Innovation: Is R&D Earning its Investment?*, DELOITTE 1, 9 (2011), https://www.deloitte.com/assets/Dcom-Poland/Local%20Assets/Documents/Raporty,%20badania,%20rankingi/pl_raport_chemiczny_2011.pdf.

and that offer no significant improvement in treatment.¹⁰ Even though these newer drugs may not be a substantial improvement over older drugs, patent protection permits companies to charge a premium.¹¹ In one extreme example, Sanofi introduced a new cancer drug at \$11,000 a month—over twice the cost of existing drugs—even though it was not more effective.¹² Given the financial constraints many nations face, providing patents on drugs of minimal therapeutic value seems especially questionable.

Although there is generally no market for expensive new products that are not a significant improvement, the pharmaceutical market is unique. Consumers generally lack expertise to assess whether new products are better, such that advertising can be particularly influential. In addition, although most patented drugs are drugs prescribed by doctors with expertise to understand the value of these drugs, their ability to effectively determine whether a new drug is worthwhile is still compromised. Doctors generally make decisions based on data by self-interested drug companies since no independent data concerning the value of a new drug generally exists. At times, the company data concerning an “improvement” is later discovered to be unsubstantiated.¹³ In addition, doctors are also influenced by

10. See, e.g., EUR. COMM’N DIR.-GEN. FOR COMPETITION, PHARMACEUTICAL SECTOR INQUIRY: FINAL REPORT 33, 351–65 (2009) [hereinafter EC PHARMACEUTICAL SECTOR INQUIRY]; JOHN R. THOMAS, PATENT EVERGREENING: ISSUES IN INNOVATION AND COMPETITION, CONGRESSIONAL RESEARCH SERVICE (2009); CHAN PARK ET AL., USING LAW TO ACCELERATE TREATMENT ACCESS IN SOUTH AFRICA: AN ANALYSIS OF PATENT, COMPETITION AND MEDICINES LAW, UNITED NATIONS DEVELOPMENT PROGRAMME 42 (2013). Although incremental innovation is common in all industries, what happens in the pharmaceutical industry is likely unique. The practice of patenting follow-on drugs is done in combination with substantial and usually successful marketing to consumers and doctors to ‘switch’ to a newly patented drug immediately before expiration of the patent for the first drug. E.g., EC PHARMACEUTICAL SECTOR INQUIRY, *supra* note 10, at 351–52. Although manufacturers of other new products, such as consumer electronics, may also market new products heavily, consumers generally cannot effectively evaluate whether a new drug is worthwhile; in addition, doctors often decide on drugs, yet are vulnerable to advertising even though they may not be conscious of this.

11. Not only can patent holders charge a premium, but they can increase the premium at any time during the patent term. This seems to be more of a problem in recent years. E.g., Robert Langreth, *Big Pharma’s Favorite Prescription: Higher Prices*, BLOOMBERG BUSINESSWEEK, May 8, 2014.

12. E.g., Peter B. Bach et al., *In Cancer Care, Cost Matters*, N.Y. TIMES, Oct. 14, 2012, at A25; see also Stephen S. Hall, *The Cost of Living*, N.Y. MAG. (Oct. 20, 2013), <http://nymag.com/news/features/cancer-drugs-2013-10/>.

13. E.g., Cynthia M. Ho, *Drugged Out: How Cognitive Bias Hurts Drug Innovation*, 51 SAN DIEGO L. REV. 419, 502–05 (2014).

pharmaceutical marketing.¹⁴ Given these issues, permitting minimally improved drugs to be patentable is problematic. Accordingly, countries should have the ability to tailor their patent laws to avoid expending resources on drugs of questionable value. Though they technically retain the right to modify domestic patent laws consistent with TRIPS, countries may consider the threat of investor-state litigation pursuant to investor-state arbitration agreements to essentially eliminate this option.

Although Eli Lilly's claim is one of first impression, it actually contributes to a broader trend. Companies are increasingly challenging domestic decisions pursuant to bilateral and multilateral agreements that provide protection to foreign investors and permit them to bring investor-state disputes. For example, investors filed only one dispute in 1982, over fifty new cases in 2012, and today there are currently five hundred claims pending in over fifty countries.¹⁵ Many have noted that such provisions threaten the ability of nations to regulate in areas of traditional domestic competence such as environmental law and public health¹⁶ because the financial stakes are often substantial—there are currently over one hundred pending actions worth more than \$1 billion each.¹⁷ Against this backdrop, Eli Lilly's suit is arguably the latest expansion of investor claims that challenge domestic laws. Moreover, many more suits challenging domestic intellectual property decisions may follow.¹⁸ Even before Eli Lilly brought suit, the multinational law firm Jones Day published a report proclaiming that investment treaty protection was “a new way forward” for multinational pharmaceutical

14. Although physicians recognize this problem, they seem to always assume that this is only a problem for *other* physicians, but that they themselves are immune from this phenomenon. *E.g., id.* at 503–04.

15. U.N. Conference on Trade & Dev., *Recent Developments in Investor-State Dispute Settlement (ISDS): Updated for the Multilateral Dialogue on Investment*, at 2–3, U.N. Doc. UNCTAD/WEB/DIAE/PCB/2013/3/REV (May 28–29, 2013) [hereinafter UNCTAD, *Recent Developments*].

16. *See, e.g.,* Stephen J. Byrnes, *Balancing Investor Rights and Environmental Protection in Investor-State Dispute Settlement Under CAFTA: Lessons from the NAFTA Legitimacy Crisis*, 8 U.C. DAVIS BUS. L.J. 102, 109–112 (2007); Samrat Ganguly, Note, *The Investor-State Dispute Mechanism (ISDM) and a Sovereign's Power to Protect Public Health*, 38 COLUM. J. TRANSNAT'L L. 113, 116, 123–26 (1999); Jane Kelsey & Lori Wallach, “Investor-State” Disputes in Trade Pact Threaten Fundamental Principles of National Judicial Systems, PUB. CITIZEN 3 (Apr. 2012), <http://www.citizen.org/documents/isds-domestic-legal-process-background-brief.pdf>.

17. *Arbitration Scorecard 2013*, AM. LAW. (June 24, 2013), <http://www.americanlawyer.com/id=1202608198051/Arbitration-Scorecard-2013>; *see also* Shawn Donnan, *Disputes Clause Heaps Pressure on Trade Deal*, FIN. TIMES (LONDON), Mar. 11, 2014 (noting that use has soared in recent years).

18. *E.g.,* Brian King & Viren Mascarenhas, *Investment Treaty Protection for IP Rights, PATENT, TRADEMARK & COPYRIGHT L. DAILY*, Aug. 5, 2013; Sherman Kahn, *Will Patents be the Next Wave in Investor-State Arbitration?*, 7 N.Y. DISP. RESOL. LAW. 53 (2014).

companies to address an “assault” against their patents in the developing world.¹⁹

The increase in investor-state disputes and expansive claims has attracted increased attention and criticism. Although scholars have criticized such actions for years, the popular press and policymakers are now also highlighting such challenges as problematic for interfering with traditional government regulation.²⁰ Moreover, critics of investor-state disputes include developed as well as developing countries—Germany recently expressed concern even though it is a party to one hundred existing agreements that provide this remedy.²¹ Commentators also have expressed concern about including such investment provisions in pending agreements. Even the Cato

19. BAIJU S. VASANI ET AL., JONES DAY, *TREATY PROTECTION FOR GLOBAL PATENTS: A RESPONSE TO A GROWING PROBLEM FOR MULTINATIONAL PHARMACEUTICAL COMPANIES* 1–2 (2012) [hereinafter JONES DAY].

20. E.g., Barrie McKenna, *Canada Must Learn From NAFTA Legal Battles*, GLOBE & MAIL (Nov. 24, 2013), <http://www.theglobeandmail.com/report-on-business/economy/canada-must-learn-from-nafta-legal-battles/article15579209/>; Elizabeth Olson, *Growth in Global Disputes Brings Big Paychecks for Law Firms*, N.Y. TIMES, Aug. 26, 2013, at B1; Sabrina Tavernise, *Tobacco Firms' Strategy Limits Poorer Nations' Smoking Laws*, N.Y. TIMES, Dec. 13, 2013, at A1; James Traub, *Trading Up: Why a Massive Trade Deal is the Key to the Asia Pivot—and to America's Future*, FOREIGN POLICY (Feb. 21, 2014), http://www.foreignpolicy.com/articles/2014/02/21/trading_up/; Elizabeth Warren & Rosa De Lauro, *Who is Writing the TPP?*, BOSTON GLOBE (May 11, 2015), <http://www.bostonglobe.com/opinion/2015/05/11/elizabeth-warren-and-rosa-de-lauro-who-writing-tpp/2FQZAV6uz9GGQI6pe3cd0K/story.html>; see also *infra* notes 71–72, 75–86 and accompanying text (describing criticisms by scholars and policymakers).

21. E.g., Peter Clark, *Germany Throws Down Gauntlet in CETA Investor-State Negotiations*, IPOLITICS (July 27, 2014), <http://www.ipolitics.ca/2014/07/27/germany-throws-down-gauntlet-in-ceta-investor-state-negotiations/>. In addition, Australia took the bold position in 2011 that it would not include such provisions in future trade agreements. DEP'T OF FOREIGN AFFAIRS AND TRADE, AUSTRAL. GOV'T, GILLARD GOVERNMENT TRADE POLICY STATEMENT: TRADING OUR WAY TO MORE JOBS AND PROSPERITY (2011); see also *Australia to Reject Investor-State Dispute Resolution in TPPA*, INV. TREATY NEWS (Apr. 13, 2012), <http://www.iisd.org/itn/2012/04/13/news-in-brief-7/>. Its current position is that it will consider such provisions on a case-by-case basis. See, e.g., Mike Secombe, *Big Tobacco's Plan to Stub Out Plain Packaging*, SATURDAY PAPER (Mar. 8, 2014), <https://www.thesaturdaypaper.com.au/opinion/topic/2014/03/08/big-tobaccos-plan-stub-out-plain-packaging/1394197200/>. However, academics suggest that Australia should not include such provisions. E.g., MATTHEW RIMMER, *TROJAN HORSE CLAUSES: INVESTOR-STATE DISPUTE SETTLEMENT* (2014), available at http://works.bepress.com/matthew_rimmer/178 (submission to the Australian Parliament). Also, some countries, such as South Africa and Indonesia, have terminated or failed to renew agreements with such provisions. E.g., U.N. Conference on Trade & Dev., *International Investment Policymaking in Transition: Challenges and Opportunities of Treaty Renewal*, at 2–3, U.N. Doc. UNCTAD/WEB/DIAE/PCB/2013/9 (June 2013); Ben Bland & Shawn Donnan, *Indonesia to Terminate More than 60 Bilateral Investment Treaties*, FIN. TIMES (Mar. 26, 2014), <http://www.ft.com/cms/s/0/3755c1b2-b4e2-11e3-af92-00144feabdc0.html>.

Institute, which usually promotes corporate interests, has suggested that a current U.S. trade initiative involving a dozen countries, the Transatlantic Trade and Investment Partnership, should not include an investment chapter not only because of concerns about domestic sovereignty, but also because investment chapters are “ripe for exploitation by creative lawyers.”²² Strong public criticism has stalled—or threatens to stall—discussions of the Trans-Pacific Partnership, as well as two bilateral agreements involving the European Union.²³ Although both the United States and the European Union at one point defended investment chapters against critics, in the past year, the European Union has stopped doing so.²⁴ The European Union is now engaging in public consultations as well as proposing modified language with the hope of minimizing concern.²⁵

Eli Lilly’s suit brings to light the problems with permitting an expansive interpretation of investment chapters to cover intellectual property “rights” that have been denied or cancelled under domestic law and are consistent with international law, such as TRIPS.²⁶ In particular, Eli Lilly’s suit highlights

22. Daniel J. Ikenson, *A Compromise to Advance the Trade Agenda: Purge Negotiations of Investor-State Dispute Settlement*, CATO INST. (Mar. 4, 2014), <http://www.cato.org/publications/free-trade-bulletin/compromise-advance-trade-agenda-purge-negotiations-investor-state>; see also *Investor-State Arbitration System Needs ‘Complete Overhaul,’* BILATERALS.ORG (May 12, 2014), <http://www.bilaterals.org/?investor-state-arbitration-system> (a prominent international lawyer who has himself served as counsel in several cases also suggested a problem with the system in delivering the keynote address to the annual Investment Treaty Arbitration Counsel).

23. E.g., Clark, *supra* note 21; Shawn Donnan & James Politi, *Official Warns EU-US Trade Deal at Risk over Investor Cases*, FIN. TIMES (Mar. 27, 2014), <http://www.ft.com/intl/cms/s/0/58ee2da0-b5dc-11e3-b40e-00144feabdc0.html>; Matthew Schwel, *EU, Canada Fail to Close CETA: Stuck Over Issue Related to Eli Lilly Case*, INSIDE U.S. TRADE, May 8, 2014; George Monbiot, Comment, *This US Trade Deal is a Full-Frontal Assault on Democracy*, GUARDIAN, Nov. 5, 2013; Brent Patterson, *Fast-Track authority stalls the Trans-Pacific Partnership*, RABBLE (May 25, 2015), <http://rabble.ca/blogs/bloggers/brent-patterson/2015/05/fast-track-authority-stalls-trans-pacific-partnership/3>.

24. E.g., EUR. COMM’N, INCORRECT CLAIMS ABOUT INVESTOR-STATE DISPUTE SETTLEMENT (2013); *The Facts on Investor-State Dispute Settlement: Safeguarding the Public Interest and Protecting Investors*, OFF. U.S. TRADE REPRESENTATIVE (Mar. 27, 2014), <http://www.ustr.gov/about-us/press-office/blog/2014/March/Facts-Investor-State%20Dispute-Settlement-Safeguarding-Public-Interest-Protecting-Investors>; *Investor-State Dispute Settlement (ISDS), Fact Sheet*, OFF. U.S. TRADE REPRESENTATIVE (Mar. 2015), <https://ustr.gov/about-us/policy-offices/press-office/fact-sheets/2015/march/investor-state-dispute-settlement-isds/>.

25. E.g., Press Release, European Comm’n, Commission to Consult European Public on Provisions in EU-US Trade Deal on Investment and Investor-State Dispute Settlement (Jan. 21, 2014).

26. This article intentionally focuses on intellectual property rights that are denied or canceled, rather than any case where intellectual property rights are at issue. Although there are pending disputes concerning whether existing trademarks that have lost value due to

an important, yet unresolved tension: a country could be vulnerable to an investor-state dispute even if it complies with a separate international law concerning intellectual property rights. This Article aims to not only evaluate the merits of and the policy problems raised by Eli Lilly's specific suit, but also demonstrate other ways that domestic attempts to balance interests of multinational drug companies with public health might be compromised.²⁷ Eli Lilly's suit may prompt other companies to challenge not only patentability standards they disagree with, but also exceptions to patent rights, even where these exceptions are permissible under TRIPS. This would threaten recent and proposed patent laws that commentators have hailed as promoting a better balance of patent rights and access to medicine.

Moreover, the threat extends beyond the patent arena: pharmaceutical companies may also use investor-state disputes to challenge domestic regulatory laws that negatively impact their ability to sell even patented drugs. Domestic laws governing clinical data associated with approval of new drugs are likely an issue. Typically, nations require companies to provide clinical data to establish that a new drug is safe and effective before approving the drug for sale.²⁸ However, companies that seek to sell a new drug are the ones that create the data and often overstate benefits.²⁹ To address this problem, the European Union has an important new law that aims to ensure that doctors and patients have full information about underlying data of approved drugs, such that they are less likely to be susceptible to overstated marketing

their inability to be used pursuant to plain packaging laws for tobacco, these raise additional issues beyond the scope of a single article. In addition, this is an important issue that has thus far been overlooked by scholars, even though some scholars have previously discussed the pending trademark cases, how compulsory licenses could be challenged as expropriations, and whether intellectual property rights in general constitute investments. *E.g.*, Christopher S. Gibson, *Latent Grounds in Investor-State Arbitration: Do International Investment Agreements Provide New Means to Enforce Intellectual Property Rights?*, in *YEARBOOK ON INTERNATIONAL INVESTMENT LAW & POLICY 2009–2010*, at 397 (Karl P. Sauvant ed., 2010); Bryan Mercurio, *Awakening the Sleeping Giant: Intellectual Property Rights in International Investment Agreements*, 15 *J. INT'L ECON. L.* 871 (2012).

27. This article focuses on issues that have not received much academic attention and also seem currently ignored in negotiations. However, there are other public health issues that have been raised as concerns. *E.g.*, Shawn Donnan, *EU Pledges to Protect NHS in US Trade Talks*, *FIN. TIMES* (July 10, 2014), <http://www.ft.com/intl/cms/s/0/eb1e1102-085e-11e4-9380-00144feab7de.html> (asserting that decisions of the National Health Service and other European public health programmes will be exempt from investor-state challenges); Press Release, HAI Europe et al., 'Investor-to-State Dispute Settlement' in EU-US Trade Deal Risks Access to Affordable Medicines (July 14, 2014), *available at* <http://haieurope.org/wp-content/uploads/2014/07/Press-release-Joint-response-ISDS-in-TTIP.pdf>.

28. *See infra* note 316 and accompanying text.

29. *See infra* Subsection III.B.2.b); Subsection IV.A.3.b).

claims.³⁰ Although many commentators heralded this law as important, it is vulnerable to challenge by companies, as later explained.³¹

This Article argues that it is unprincipled to permit companies to use investor-state arbitration to challenge domestic decisions when they have no valid intellectual property rights. Eli Lilly, for example, is asserting that it is entitled to compensation because two of its initially issued patent were later invalidated for failing to satisfy mandatory patent law requirements. However, its invalidated patents amount to rights that never existed in the first place. It would be more analogous to a void contract, which purports to give rights but was never enforceable. Moreover, unlike most other types of property, the very existence of intellectual property is only justified if it promotes desired policy. Most recognize a utilitarian justification for intellectual property—to promote innovation while at the same time recognizing and balancing competing policy considerations, such as access for researchers, users, and consumers.³² Although some believe that intellectual property rights are inherently beneficial in promoting innovation and attracting foreign direct investment, data does not clearly support this claim.³³ Since such rights inherently reduce access and increase costs, countries should view the policy justification for providing such rights skeptically. In addition, although there are international agreements that require countries to provide equal treatment to domestic and foreign intellectual property right holders, all such agreements assume that each country has some discretion to decide how to balance intellectual property rights against other interests.

This Article also argues that permitting companies to challenge domestic decisions regarding intellectual property through investor-state disputes is problematic because they disrupt internationally agreed norms under TRIPS, and also because the historical justifications for protecting foreign investors do not apply. TRIPS permits member countries discretion in shaping intellectual property rights to advance policy goals other than promoting

30. *See infra* Subsection IV.A.3.b).

31. *See, e.g.*, HAI EUROPE, PROTECTING CITIZENS' HEALTH: TRANSPARENCY OF CLINICAL TRIAL DATA ON MEDICINES IN THE EU (2013).

32. Of course, there are competing views on whether inventors and artists have a “natural” right to intellectual property, as well as whether these rights are protected as human rights instruments, such as the International Covenant on Economic, Social and Cultural Rights. *See* International Covenant on Economic, Social and Cultural Rights, G.A. Res. 2200A (XXI), art. 15.1(b), U.N. Doc. A/RES/21/2200 (Dec. 16, 1966). This Article agrees with the majority of commentators who reject this view. Historical and policy justifications of intellectual property rights support the majority view—patents have always served to further the societal goals of promoting and disseminating technological innovation.

33. *See infra* Subsection III.A.2.c).

innovation, including public health.³⁴ A country that arguably complies with TRIPS should not be subject to an investor-state challenge that could disrupt TRIPS norms, as well as result in a judgment inconsistent with a decision from the World Trade Organization (“WTO”)—the international body that handles TRIPS disputes. Moreover, permitting such challenges would not further the historical justification of protecting investments in international investment agreements—to promote investment from foreign companies. Unlike most other types of investments where increased protection induces foreign investment, companies do not invest in countries solely, or even primarily due to the strength of intellectual property protection; in addition, whether or not a company decides to seek more patent protection is unlikely to benefit the host country unlike the typical scenario where stronger protection for foreign companies results in direct benefit to the host country. Eli Lilly’s case uniquely threatens a country’s freedom to legislate under a separate international agreement. In contrast, prior instances of investor-state disputes that arguably relate to another international agreement have not resulted in a clear conflict pursuant to which resolution of the investor’s claim would violate the integrity of a separate international agreement. Often this is because the other international agreement only is aspirational.³⁵

This Article also provides a detailed analysis of Eli Lilly’s central claims—that Canada “expropriated” Eli Lilly’s investment in its patents, and that Canada failed to provide “fair and equitable treatment” to Eli Lilly’s investments.³⁶ The Article seeks to show why Eli Lilly should not recover on

34. TRIPS, *supra* note 3, art. 1.1; World Trade Organization, Ministerial Declaration of 14 November 2001, WT/MIN(01)/DEC/2, ¶¶ 4–5 (2002) [hereinafter Doha Public Health Declaration]; *see also infra* Subsection II.B.1 (discussing flexibility under TRIPS).

35. *See infra* notes 138–42 and accompanying text (discussing nonbinding WHO framework convention as well as ambiguous human rights norms).

36. Although Eli Lilly originally asserted that Canada violated the nondiscrimination agreement, that provision will not be addressed because it lacks any merit; indeed, Eli Lilly’s notice of arbitration dropped this ground. In particular, national treatment simply requires a country to treat foreign investors “no less favorabl[y]” than its domestic investors. North American Free Trade Agreement, U.S.-Can.-Mex., ch. 11, art. 1102, Dec. 17, 1992, 32 I.L.M. 605, 639 [hereinafter NAFTA]. Eli Lilly made two claims that seem unmoored to this standard. Eli Lilly claimed that Canadian law disadvantages foreign nationals with requirements “not required by the foreign applicants’ own national jurisdictions.” *Eli Lilly & Co. v. Canada*, Notice of Intent to Submit a Claim to Arbitration Under NAFTA Chapter Eleven, ¶ 106 (NAFTA/UNCITRAL Arb. Trib. Nov. 7, 2012). However, nondiscrimination does not guarantee an investor laws identical to its home state. In addition, Eli Lilly claimed that it is treated less favorably than domestic generic competitors that can benefit from making the now invalidated patented drugs. *Id.* ¶ 107. However, nondiscrimination is only about comparing similarly situated entities, and generic pharmaceutical companies are not similar—they have an entirely different business model, such that this claim is illogical.

any of its claims, but nonetheless explains which issues are most vulnerable for Canada. As the final Part of this Article highlights, recognizing potential problems that may arise from the decision in Eli Lilly's case is important to understanding how to properly cabin such claims in future cases.

This Article proceeds in three parts: Part II provides a background of relevant domestic and international law and policy, as well as an overview of Eli Lilly's claims to contextualize Eli Lilly's investor-state dispute. Part III then turns to the specifics of Eli Lilly's claims against Canada and explains why an arbitration tribunal should reject Eli Lilly's claims. Part IV goes beyond the specifics of Eli Lilly's claims to explain other TRIPS-consistent domestic laws that are in danger of subsequent investor-state disputes. This Part concludes with specific proposals for preserving countries' existing policy space under TRIPS in pending and future agreements concerning investor protection.

I. BACKGROUND

A. DOMESTIC PATENT LAW AND POLICY

To understand Eli Lilly's claims against Canada, some background on patent law is important. A patent is a legal document granted by a country to the creator of an invention that provides the commercially valuable ability to exclude others from making or selling the patented invention within the boundaries of the patent-granting country.³⁷ A patent is fundamentally a lever to promote social policy, offering a reward to induce inventors to disclose information to society so that others can learn from and build upon that innovation.³⁸ Because most inventions build upon prior inventions, encouraging inventors to share their knowledge is socially valuable, even if there is a temporary cost of higher prices during the period of patent protection. A patented drug is generally expensive because the patent owner can exclude all others from making the identical drug during the patent term; therefore the patent owner can and generally does charge a substantial premium.

The social harm of higher prices on patented goods is mediated by a limited term of patent protection, as well as patentability requirements. Patents generally award inventors a term of protection of no more than twenty years to minimize the period during which consumers must pay

37. *E.g.*, 35 U.S.C. § 271 (2012); TRIPS, *supra* note 3, art. 28.

38. *See* *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 481 (1974); *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*, 489 U.S. 141, 164–65 (1989).

patent-inflated prices.³⁹ In addition, the patentability standards aim to ensure that patents are only granted on more socially valuable inventions so that the price premium only applies where deserved.

There are two basic types of requirements that patent applications must satisfy. First, the invention must meet certain patentability requirements, typically that it is: patentable subject matter, useful (or that it has “utility”), new, and not obvious.⁴⁰ For example, an invention that has no use at all, or only a “throw away” use, such as being used as a paperweight, would not deserve a patent. However, nations differ on the appropriate standard for utility. The United States, for example, has one of the broadest interpretations of usefulness; an invention qualifies as useful by providing simple entertainment, rather than a commercial use, and may also be useful even if its use promotes deceptive or even illegal activity in some circumstances.⁴¹ Most European countries, on the other hand, bar patents on inventions that violate morality.⁴² Second, the application itself must meet certain disclosure requirements—it must fully describe the invention and enable others to properly make and use the disclosed invention.⁴³

A patent application must satisfy these patentability requirements at the time of filing to foster fundamental patent policy goals. In particular, patentability requirements prevent applicants from filing claims for inventions that they have not fully developed to avoid unduly rewarding speculative claims, which could bar research and impose costs on the public. Notably, although Canada has a unique interpretation of the utility requirement under the “promise doctrine,” its interpretation in fact supports this fundamental patent policy. In cases where an application promises a certain use (but not if there is no such promise), an invention is only considered useful if it provides an adequate disclosure for that promise. The Canadian Supreme Court has recognized costs of patent protection to be particularly important in the area of pharmaceuticals, noting that

39. *E.g.*, 35 U.S.C. § 154(a)(2) (2012); TRIPS, *supra* note 3, art. 33.

40. *E.g.*, 35 U.S.C. §§ 101–103 (2012); TRIPS, *supra* note 3, art. 27(1).

41. *E.g.*, *Juicy Whip, Inc. v. Orange Bang, Inc.*, 185 F.3d 1364, 1368 (Fed. Cir. 1999). In addition, there are patents on inventions that may be illegal such as radar detectors and gambling devices.

42. Convention on the Grant of European Patents art. 53(a), Oct. 5, 1973, 1065 U.N.T.S. 199 (barring patents on inventions “the publication or exploitation of which would be contrary to ‘ordre public’ or morality”) [hereinafter European Patent Convention]; *see also* Cynthia M. Ho, *Splicing Morality and Patent Law: Issues Arising from Mixing Mice and Men*, 2 WASH. U. J.L. & POL’Y 247 (2000) (explaining cases interpreting this doctrine).

43. *E.g.*, 35 U.S.C. § 112 (2012); TRIPS, *supra* note 3, art. 29.

[w]ere the law to be otherwise, major pharmaceutical companies could . . . patent whole stables of chemical compounds for all sorts of desirable but unrealized purposes in a shot-gun approach hoping that, as in a lottery, a certain percentage of compounds will serendipitously turn out to be useful for the purposes claimed.⁴⁴

Although Canada's interpretation of what is useful is different than the interpretations of other countries, countries widely recognize the purpose espoused by the Canadian Supreme Court. The U.S. Supreme Court has similarly stated that a patent is not a "hunting license" in interpreting the utility requirement to bar a patent on a new method of making a known steroid when the inventors had no idea what function the steroid served, other than as a tool for further research.⁴⁵

Like all types of intellectual property, patents are granted by individual nations and are territorially limited. Because a global patent does not exist, an inventor must seek patent protection in individual countries. For example, a U.S. patent gives its owner rights against others in the U.S., but does not provide any protection in other countries. To obtain protection in other countries, the inventor would have to apply for and obtain patent protection in the desired countries. However, an identical patent application filed in different countries will not result in identical patent rights, or even patentability. This is because each country has its own patentability standards. Contrary to Eli Lilly's suggestions, courts in different countries can and often do apply different domestic standards to determine whether the same invention is patentable; courts often note that other nations' decisions regarding similar, or even identical patent applications are not relevant to a domestic court's evaluation of patentability.⁴⁶

B. THE INTERNATIONAL LANDSCAPE: IP AND FOREIGN INVESTMENTS

Before addressing the specifics of Eli Lilly's claims, it may help to provide a broader context of the international landscape. As noted in the introduction, Eli Lilly's case lies at the unique intersection of separate

44. *Apotex Inc. v. Wellcome Foundation, Ltd.*, [2002] 4 S.C.R. 153, para. 80 (Can.).

45. *Brenner v. Manson*, 383 U.S. 519, 535–36 (1966).

46. *E.g.*, *Canada (Att'y Gen.) v. Amazon.com, Inc.*, [2011] F.C.A. 328, para. 16 (Can.) ("[I]t would not be helpful . . . to explain the results of Amazon's patent applications in other jurisdictions. . . . [E]very jurisdiction has its own patent laws and administrative practices, and they are inconsistent with one another in important respects."); *Conor Medsystems Inc. v. Angiotech Pharm. Inc.*, [2008] UKHL 49, ¶ 3 (stating that "it is inevitable that [different courts] will occasionally give inconsistent decisions about the same patent"); *Apotex*, [2002] 4 S.C.R. 153, para. 40 (Can.) ("[G]iven the differences in our respective patent laws, the outcome of the U.S. litigation on this patent is of limited interest here.").

international agreements that govern intellectual property and foreign investments.⁴⁷ To better understand Eli Lilly's case, this Section first considers the traditional territorial scope of intellectual property rights, as well as how such rights are treated under international intellectual property agreements such as TRIPS. Finally, this Part provides an overview of bilateral and multilateral agreements that protect foreign direct investment.

1. *Intellectual Property: Territoriality and Treaties*

There is a long history of recognizing national boundaries of patents in the global arena. The Paris Convention was one of the earliest international agreements relating to patents (established in 1883), yet it did not require any member countries to grant patents—this was considered a domestic choice.⁴⁸ The agreement only provided rules to ensure fairness to domestic and foreign applicants, and facilitate the process of obtaining patents in multiple countries.⁴⁹ Indeed, a provision on the independence of patents specifically clarified that the grant or denial of a patent on an invention in one country does not impact the decision of another country.⁵⁰ This rule inherently recognized and reinforced that patents are tools of domestic discretion and that patent rights, if any, are restricted to the territory of the patent-granting nation. Accordingly, countries not only differed on whether or not to completely ban patents, but also on the grounds for granting patents; nations could decide to grant patents only on processes, or only in some fields of technology.

The second significant international agreement relating to patents is the Patent Cooperation Treaty ("PCT"). This agreement, concluded in 1970, provided a mechanism for inventors to more easily obtain patents in multiple countries with a uniform PCT application that could be examined in individual member states. Importantly, the PCT still firmly recognized territorial limits and domestic sovereignty. It simply made the process of filing for patents in multiple countries easier with a single application. For example, a PCT application establishes a priority filing date in all member

47. In some situations, including Eli Lilly's, there is an international agreement that simultaneously governs both intellectual property and foreign investments in separate chapters of the agreement. However, the issues are the same as if there were independent agreements.

48. Paris Convention for the Protection of Industrial Property, Mar. 20, 1883, as last revised at the Stockholm Revision Conference, July 14, 1967, 21 U.S.T. 1583, 828 U.N.T.S. 305 [hereinafter Paris Convention].

49. For example, a nation that provided patents could not treat patent applications of its own nationals better than those of foreign applicants. Paris Convention, *supra* note 48, arts. 2–3 (national treatment).

50. Paris Convention, *supra* note 48, art. 4*bis*.

countries, but each country where the applicant ultimately files evaluates the patent according to its own patent standards.⁵¹

Today, most nations of the world are members of the WTO and, as such, must provide patent and other intellectual property rights pursuant to TRIPS. This landmark agreement, concluded in 1994, has dramatically changed the ability of nations to exercise full control over the decision to grant patent rights. Member countries are required to grant patents to all “inventions,” that meet certain standards of patentability.

Even though TRIPS now requires many countries to provide patents, because the agreement sets minimum, rather than uniform standards, it explicitly contemplates diversity in domestic laws.⁵² Although TRIPS notably requires some nations to provide patents for the first time, it gives states substantial flexibility in how to do so. For example, although TRIPS requires nations to grant patents on “inventions” that meet patentability standards, it does not define what constitutes an invention.⁵³ Accordingly, nations can properly exclude software, for example, from patentability if they do not consider software to be an invention.⁵⁴ Similarly, although TRIPS requires nations to provide patents on inventions that are useful, new, and nonobvious, it does not define any of these terms.⁵⁵ Before TRIPS, member states had different laws about some of these terms. The lack of inclusion of any specific definitions permits nations to provide their own definitions.⁵⁶ A failed attempt to create a new patent law treaty with uniform patent standards after TRIPS underscores that nations understood TRIPS not to

51. Patent Cooperation Treaty, June 19, 1970, 28 U.S.T. 7645, 1160 U.N.T.S. 231 [hereinafter Patent Cooperation Treaty].

52. TRIPS, *supra* note 3, art. 1(1).

53. *Id.* at art. 27.

54. *Id.*

55. *Id.*

56. *E.g.*, Carlos M. Correa, *Patent Rights*, in INTELLECTUAL PROPERTY AND INTERNATIONAL TRADE: THE TRIPS AGREEMENT 189, 198–200 (2d ed. 1998); PARK, *supra* note 10, at 39; WORLD HEALTH ORG. ET AL., PROMOTING ACCESS TO MEDICAL TECHNOLOGIES AND INNOVATION: INTERSECTIONS BETWEEN PUBLIC HEALTH, INTELLECTUAL PROPERTY AND TRADE 57 (2012); Michael N. Meller, *Principles of Patentability and Some Other Basics for a Global Patent System*, 83 J. PAT. & TRADEMARK OFF. SOC'Y 359, 359 (2001); E. Richard Gold & Michael Shortt, *The Promise of the Patent in Canada and Around the World*, 30 CANADIAN INTELL. PROP. REV. 35, 58–59 (2014); Jerome H. Reichman & Rochelle Cooper Dreyfuss, *Harmonization Without Consensus: Critical Reflections on Drafting a Substantive Patent Law Treaty*, 57 DUKE L.J. 85, 89 (2007); Expert Report of Daniel Gervais Accompanying Counter Memorial of Canada, ¶¶ 22–54.

impose such uniformity.⁵⁷ The attempt to create uniform standards shows that countries recognized that there were no uniform standards of patentability under existing agreements, such as TRIPS.

Accordingly, although member countries must provide some protection to drug patents, they can define TRIPS patentability criteria to minimize harm to social policies beyond innovation. For example, a country could consider a newly discovered use of a known compound to neither be an “invention” nor “new” because TRIPS does not define these key terms. Although the use itself may be new, arguably that was inherent in the existence of the compound, such that society does not benefit from granting a patent on this new use. Countries might also consider such a use as simply falling below the threshold inventiveness. In fact, India’s recent patent amendments take this approach in barring from patentability new uses of known compounds.⁵⁸

Countries may also decide how to satisfy the TRIPS patentability requirement that an invention be “useful.”⁵⁹ TRIPS article 27 actually states that countries must provide patents on inventions that are “capable of industrial application,” but clarifies in a footnote that this term is meant to be synonymous with “useful.” These terms refer to somewhat similar, but slightly different patentability requirements that existed in the laws of countries at the time TRIPS was negotiated. As noted by one scholar, “the deliberate inclusion of these two alternatives precludes any inference that the draftsmen of TRIPS intended to incorporate by reference or implication any single existing standard of patentability.”⁶⁰

Indeed, these differences have always been permissible. Prior to the conclusion of TRIPS in 1994, the Paris Convention simply focused on ensuring fairness to domestic and foreign applicants if a country decided to provide patents and the PCT similarly facilitated the process of obtaining

57. Reichman & Dreyfuss, *supra* note 56, at 89–90. This effort was abandoned in 2006. See *Draft Substantive Patent Law Treaty*, WIPO, http://www.wipo.int/patent-law/en/draft_splt.htm (last visited May 9, 2015).

58. The Patents (Amendment) Act, 2005, No. 15, § 3(d), Acts of Parliament, 2005 (India).

59. TRIPS, *supra* note 3, art. 27(1).

60. Christopher Wadlow, *Utility and Industrial Applicability*, in *PATENT LAW AND THEORY: A HANDBOOK OF CONTEMPORARY RESEARCH* 355, 356 (Toshiko Takenaka, ed., 2008).

patents in multiple countries.⁶¹ These agreements remain in force today, and the principle of territoriality remains valid.⁶²

2. *International Agreements Governing Foreign Direct Investments and Providing Investor-State Disputes*

There are over 3000 international agreements that provide foreign investors substantive rights to protect their investment, as well as a mechanism to protect those rights outside of domestic courts.⁶³ These agreements are either bilateral investment agreements, or free trade agreements, such as the North American Free Trade Agreement (“NAFTA”), with an investor chapter aimed at promoting foreign investment.⁶⁴ Although a free trade agreement may include a chapter on both intellectual property rights and foreign investor rights, these chapters operate independently of each other.

Most agreements protecting foreign investors provide a similar set of substantive rights. These rights build upon prior international rights and provide more clarity than prior principles of customary international law. Typically, they provide investors a guarantee of compensation for any expropriation of investments, freedom from unreasonable or discriminatory measures, fair and equitable treatment, and guarantee foreign investors that they will not be treated less favorably than domestic ones. All of these rights

61. Paris Convention, *supra* note 48, arts. 2–3; PCT, *supra* note 51 (streamlining process to file patents in multiple countries).

62. In addition, TRIPS clarifies that it is consistent with the Paris Convention. TRIPS, *supra* note 3, art. 2.

63. *E.g.*, UNCTAD, INVESTOR-STATE DISPUTES: PREVENTION AND ALTERNATIVES TO ARBITRATION 3–4 (2010), available at http://unctad.org/en/docs/diaeia200911_en.pdf; INT’L INST. FOR ENVIRON. & DEV. (IIED), INVESTMENT TREATIES AND SUSTAINABLE DEVELOPMENT: INVESTOR-STATE ARBITRATION 1 (2014), available at <http://pubs.iied.org/pdfs/17241IIED.pdf>.

64. Although scholars have questioned whether any of these agreements in fact impact decisions to invest, that is nonetheless the goal. *E.g.*, JESWALD W. SALACUSE, THE LAW OF INVESTMENT TREATIES 121 (2009) (noting conflicting studies); Todd Allee & Clint Peinhardt, *Delegating Differences: Bilateral Investment Treaties and Bargaining Over Dispute Resolution Provisions*, 54 INT’L STUD. Q. 1 (2010); Susan D. Franck, *Foreign Direct Investment, Investment Treaty Arbitration, and the Rule of Law*, 19 PAC. MCGEORGE GLOBAL BUS. & DEV. L.J. 337 (2007); Jeswald W. Salacuse & Nicholas P. Sullivan, *Do BITs Really Work?: An Evaluation of Bilateral Investment Treaties and Their Grand Bargain*, 46 HARV. INT’L L.J. 67, 75–79 (2005); Tom Ginsburg, *International Substitutes for Domestic Institutions: Bilateral Investment Treaties and Governance*, 25 INT’L. REV. L. & ECON. 107, 108 (2005). *But see* Todd Allee & Clint Peinhardt, *Contingent Credibility: The Impact of Investment Treaty Violations on Foreign Direct Investment*, 65 INT’L ORG. 401 (2011); Jennifer L. Tobin & Susan Rose-Ackerman, *When BITs Have Some Bite: The Political-Economic Environment for Bilateral Investment Treaties*, 6 REV. INT’L ORGS. 1 (2011).

help to ensure that host governments will not subject foreign investors to inappropriate risks, and consequently induce them to invest.

Investor-state arbitrations developed simultaneously with bilateral investment treaties as a means to promote and protect investment from foreign companies. These largely began after World War II when newly independent nations wanted to encourage foreign investment and assistance with developing natural resources, as well as to generally encourage foreign investment.⁶⁵ The investor-state arbitration remedy provided an important avenue for relief to investors. Although foreign investors previously might have attempted to sue the state in its own courts, those courts could be biased; alternatively, the state might be able to claim sovereign immunity. Sometimes the investor could not even directly pursue an action. In the worst-case scenario, home states used, or at least threatened to use, military force.⁶⁶ All of these options provided the investor with limited avenues of recourse, either because no suit could be brought or because a judgment was not enforceable.⁶⁷

Investor-state arbitrations address this problem through a unique process. Investors bring claims not before a domestic or international court, but a tribunal of private arbitrators, who are generally lawyers. The state is considered to have consented to this by agreeing to the treaty provision. The ability of foreign investors to arbitrate their disputes against states obviates prior hurdles to protecting investments when domestic courts either did not

65. Germany and Pakistan signed the first such agreement in 1959. United Nations Commission on Trade and Development, *Bilateral Investment Treaties in the Mid 1990s*, UNCTAD/ITE/IIT/7, Sales No. E. 98.II.D.8, 8–10 (1998). International agreements to promote investments date back as far as the late 1700s, but mostly focused on expropriation and also did not have the same type of arbitration-based remedies as current agreements. *E.g.*, RUDOLF DOLZER & CHRISTOPH SCHREUER, *PRINCIPLES OF INTERNATIONAL INVESTMENT LAW* 6–8 (2d ed. 2012).

66. *E.g.*, Susan Franck, *Development and Outcomes of Investment Treaty Arbitration*, 50 HARV. INT'L L.J. 435, 442 (2009) (noting gunboat diplomacy and war); KENNETH J. VANDEVELDE, *UNITED STATES INVESTMENT TREATIES: POLICY AND PRACTICE* 7 (1992) (describing use of military force); Ray C. Jones, *NAFTA Chapter 11 Investor-to-State Dispute Resolution: A Shield to be Embraced or a Sword to be Feared?*, 2002 BYU L. REV. 527, 529 (describing gun boat diplomacy).

67. For example, even if the ICJ ruled in favor of a company, a nation might decide not to pay the investor. The only remedy in such a case was passing a UN Security Council resolution. *E.g.*, Susan D. Franck, *The Legitimacy Crisis in Investment Treaty Arbitration: Privatizing Public International Law Through Inconsistent Decisions*, 73 FORDHAM L. REV. 1521, 1537 (2005).

recognize any claims, or refused to enforce domestic judgments in favor of foreign investors.⁶⁸

Although the global community initially praised this process as a way to promote investment flows during a more stable global economic era, more recently, scholars, countries, and citizens have criticized multiple aspects of investor-state arbitrations.⁶⁹ Some of the criticisms build on popular anti-trade or anti-corporate sentiment.⁷⁰ However, countries as well as scholars have also expressed reservation and criticism.⁷¹ A major issue is that the suits appear to improperly encroach on domestic authority and even have a chilling effect on legitimate state regulatory functions due to substantial awards, as well as legal costs of defending such cases.⁷² One example of a substantial award occurred in July 2014 against Russia for over \$50 billion.⁷³ Although a recent study noted that states win in sixty percent of the cases,

68. *E.g., id.* at 1537; Anthea Roberts, *State-to-State Investment Treaty Arbitration: A Hybrid Theory of Interdependent Rights and Shared Interpretive Authority*, 55 HARV. INT'L L.J. 1, 15–17 (2014) (discussing gunboat diplomacy as well as broader historical context).

69. For a discussion of early praise, see David P. Riesenberg, Note, *Fee Shifting in Investor-State Arbitration: Doctrine and Policy Justifying Application of the English Rule*, 60 DUKE L.J. 977, 985 (2011).

70. Citizens may not only object, but publicly protest. For example, in South Korea, there were physical fights and tear gas use. *E.g., South Korea Passes U.S. Free-Trade Agreement, Lawmaker Sets Off Tear Gas Canister in Protest*, FOX NEWS (Nov. 22, 2011), <http://www.foxnews.com/world/2011/11/22/south-korea-passes-us-free-trade-agreement-lawmaker-sets-off-tear-gas-canister/>.

71. *E.g.,* Joseph Stiglitz, Letter to Congress, May 18, 20015, available at <http://www.scribd.com/doc/265770405/Letter-to-Congress-Stiglitz-on-Trade-Deal>; Letter to Congress, April 30, 2015, available at <http://www.afj.org/wp-content/uploads/2015/03/ISDS-Letter-3.11.pdf>; GUS VAN HARTEN ET AL., PUBLIC STATEMENT ON THE INTERNATIONAL INVESTMENT REGIME ¶ 8 (Aug. 31, 2010) (statement by some academics suggesting that the process is “not a fair, independent and balanced method for the resolution of investment disputes”); *supra* note 21 and accompanying text (noting concern by countries); *but see* Hon. Charles N. Brower & Sadie Blanchard, *What's in a Meme? The Truth About Investor-State Arbitration: Why It Need Not, and Must Not, Be Repossessed by States*, 52 COLUM. J. TRANSNAT'L L. 689 (2014).

72. *E.g.,* EUR. PARLIAMENTARY RESEARCH SERV., INVESTOR-STATE DISPUTE SETTLEMENT (ISDS): STATE OF PLAY AND PROSPECTS FOR REFORM 5 (2014). The average cost of attorney fees alone has been estimated at \$5 million U.S. dollars with a mean tribunal cost of approximately \$1 million. *E.g.,* Susan D. Franck, *Using Investor-State Mediation Rules to Promote Conflict Management: An Introductory Guide*, 29 ICSID REV. 66, 77–78 (2014).

73. The Russian award is the largest award to date. See Glyn Moody, *Corporate Sovereignty Tribunal Makes \$50 Billion Award Against Russia*, TECHDIRT (Aug. 5, 2014, 12:37 AM), <https://www.techdirt.com/articles/20140801/05242228082/corporate-sovereignty-tribunal-makes-50-billion-award-against-russia.shtml>; Kavaljit Singh, *The Era of Mega-Arbitration: International Court Rules Against Russia in \$50 Billion Decision*, MADHYAM (July 30, 2014), <http://www.madhyam.org.in/the-era-of-mega-arbitration-tribunal-awards-50-billion-against-russia-in-yukos-case/>.

the average award—\$16.6 million—is nonetheless significant.⁷⁴ Although there is a huge diversity in awards, even a lower award would still be substantial for any developing country, such that a potential award, even if statistically unlikely, could have a substantial impact on domestic decisions.

A major complaint is that the system results in inconsistent decisions because there is no binding precedent,⁷⁵ tribunals interpret provisions broadly,⁷⁶ and there is no appeal system.⁷⁷ Although tribunals often rely on prior decisions⁷⁸ and awards, and counsel for parties regularly cite prior decisions, the lack of hierarchy among tribunals as compared to traditional court systems, as well as the lack of an appellate system, may result in unpredictability.⁷⁹ Some also contend that arbitrators lack the independence and impartiality of typical domestic or international tribunals.⁸⁰ Although arbitrators are clearly private parties, rather than judges, some suggest that the presumption that arbitrators will rule in favor of corporations is overstated.⁸¹ A related issue is that the proceedings and decisions may lack the same level of transparency as most judicial decisions. The rules for most proceedings do not permit interested parties to participate and do not require

74. *E.g.*, Susan D. Franck, *Using Investor-State Mediation Rules to Promote Conflict Management: An Introductory Guide*, Washington & Lee Legal Studies Paper No. 2014-13, page 14, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2389763.

75. This is true of international law in general. *E.g.*, Statute of the International Court of Justice art. 59, June 26, 1945, 59 Stat. 1055, T.S. 993, 3 Bevens 1179; MOHAMED SHAHABUDEEN, PRECEDENT IN THE WORLD COURT (1996); Gabrielle Kaufmann-Kohler, *Arbitral Precedent: Dream, Necessity or Excuse?*, 23 ARB. INT'L 357, 362–65 (2007); see also NAFTA, *supra* note 36, art. 1136(1) (“An award made by a Tribunal shall have no binding force except between the disputing parties and in respect of the particular case.”).

76. *E.g.*, Elizabeth Olson, *Growth in Global Disputes Brings Big Paychecks for Law Firms*, N.Y. TIMES (Aug. 26, 2013), <http://dealbook.nytimes.com/2013/08/26/growth-in-global-disputes-brings-big-paychecks-for-law-firms/>.

77. AUSTL. GOV'T PRODUCTIVITY COMM'N, BILATERAL AND REGIONAL TRADE AGREEMENTS: PRODUCTIVITY COMMISSION RESEARCH REPORT 273 (2010); UNCTAD, Recent Developments, *supra* note 1515, at 26.

78. Christoph Schreuer & Matthew Weiniger, *A Doctrine of Precedent?*, in THE OXFORD HANDBOOK OF INTERNATIONAL INVESTMENT LAW 1196 (Peter Muchlinski et al. eds., 2008) (considering there to be a de facto practice of precedent, even if not required).

79. See Schreuer & Weiniger, *supra* note 78, at 1191–95; Jeswald W. Salacuse, *The Emerging Global Regime for Investment*, 51 HARV. INT'L L.J. 427, 460–61 (2010).

80. *E.g.*, PIA EBERHARDT & CECILIA OLIVET, PROFITING FROM INJUSTICE: HOW LAW FIRMS, ARBITRATORS AND FINANCIERS ARE FUELLING AN INVESTMENT ARBITRATION BOOM 9 (2012); Olson, *supra* note 76; Gus Van Harten, *Arbitrator Behaviour in Asymmetrical Adjudication: An Empirical Study of Investment Treaty Arbitration*, 50 OSGOODE HALL L.J. 211, 216 (2012).

81. Gus Van Harten, *Perceived Bias in Investment Treaty Arbitration*, in THE BACKLASH AGAINST INVESTMENT ARBITRATION: PERCEPTIONS AND REALITY 433 (Michael Waibel et al. eds., 2010).

final decisions to be made public. Although there are recent rules that increase transparency, these rules only apply prospectively to new agreements, rather than to the many that already exist.⁸² These disputes are admittedly more transparent than general commercial arbitration, but there is a major distinction in that investor-state disputes involve actions against sovereign nations. Even though nations consent to this process in a manner analogous to private parties, the consent is for a broad range of claims whereas commercial arbitration generally involves claims arising from a contractual clause limited to a specific situation.

To combat these shortcomings there have been many proposals to reform the current system for investor-state disputes.⁸³ Many have suggested some type of appellate body to address the problem of inconsistent as well as expansive interpretations of identical provisions.⁸⁴ Alternatively, some suggest replacing private arbiters with an international investment court to promote impartiality and independence.⁸⁵ Other proposals do not involve drastic changes to the dispute resolution process but nonetheless aim to cabin problematic decisions. For example, some suggest requiring that claimants first exhaust domestic remedies; limiting the scope of claims; or

82. UNCITRAL Rules on Transparency in Treaty-based Investor-State Arbitration art. 1 (2014), available at <http://www.uncitral.org/pdf/english/texts/arbitration/rules-on-transparency/Rules-on-Transparency-E.pdf> (mandatory application only for disputes under treaties concluded on or after 2014).

83. An alternative approach is to renegotiate or withdraw from such agreements entirely. *E.g.*, Ben Bland & Shawn Donnan, *Indonesia to Terminate More than 60 Bilateral Investment Treaties*, FIN. TIMES, Mar. 26, 2014; Andrew Newcombe, *A Brief Comment on the "Public Statement on the International Investment Regime"*, KLUWER ARB. BLOG (Sept. 3, 2010), <http://kluwerarbitrationblog.com/blog/2010/09/03/public-statement-on-the-international-investment-regime/>; see also Gus Van Harten et al., *Public Statement on the International Investment Regime*, BILATERALS.ORG (Aug. 30, 2010), <http://bilaterals.org/?public-statement-on-the> (statement from academics asserting that "[t]here is a strong moral as well as policy case for governments to withdraw from investment treaties and to oppose investor-state arbitration, including by refusal to pay arbitration awards").

84. *E.g.*, APPEALS MECHANISMS IN INTERNATIONAL INVESTMENT DISPUTES (Karl P. Sauvant ed., 2009); Asif H. Qureshi, *An Appellate System in International Investment Arbitration?*, in THE OXFORD HANDBOOK OF INTERNATIONAL INVESTMENT LAW 1154 (Peter Muchlinski et al. eds., 2008); Mark Kantor, *ICSID SECRETARIAT, POSSIBLE IMPROVEMENTS OF THE FRAMEWORK FOR ICSID ARBITRATION*, TRANSNATIONAL DISPUTE MANAGEMENT (2004), <http://www.transnational-dispute-management.com/article.asp?key=307> (suggesting consideration of an appeals facility at ICSID). *But see* Schreuer & Weiniger, *supra* note 78, at 1203–05 (suggesting that appeals may not be the best way to provide consistent interpretation and instead suggesting preliminary rulings on questions of law instead).

85. *E.g.*, Gus Van Harten, *Commentary: A Case for an International Investment Court*, INV. TREATY NEWS, Aug. 7, 2008; U.N. Conference on Trade & Dev., *Reform of Investor-State Dispute Settlement: In Search of a Roadmap*, U.N. Doc. UNCTAD/WEB/DIAE/PCB/2013/4 (June 26, 2013).

requiring arbitrators to consider other areas of international law, such as human rights and environmental obligations.⁸⁶

C. ELI LILLY'S CASE

Eli Lilly has filed a notice of arbitration against Canada, alleging violations of NAFTA's investment chapter. Eli Lilly claims that Canadian courts improperly invalidated its patents for failing the utility requirement and challenges the Canadian "promise doctrine"—a common law interpretation of utility that applies when a patent sets forth, or is perceived to set forth, a "promise." The promise doctrine dates back to early patent law in the UK, as well as older Canadian decisions; however, Canadian courts only recently invalidated patents under this doctrine.⁸⁷ Since 2005, Canadian courts have invalidated roughly a dozen patents for failing to satisfy this doctrine.⁸⁸ Pursuant to this doctrine, a patent that promises something is only useful if it does what it "promises." If the patent does not make a promise, a scintilla of utility can establish usefulness.⁸⁹ For patents and patent applications that make a promise, whether the promise is fulfilled can either

86. *E.g.*, THOMAS McDONAGH, UNFAIR, UNSUSTAINABLE, AND UNDER THE RADAR: HOW CORPORATIONS USE GLOBAL INVESTMENT RULES TO UNDERMINE A SUSTAINABLE FUTURE 7, 15 (2013); JONAS PARELLO-PLESNER & ELENA ORTIZ DE SOLÓRZANO, A COMPREHENSIVE APPROACH TO INVESTMENT PROTECTION 2 (2013); Stephan W. Schill, *Enhancing International Investment Law's Legitimacy: Conceptual and Methodological Foundations of a New Public Law Approach*, 52 VA. J. INT'L L. 57, 69 (2011).

87. The promise doctrine has its origins in British law prior to 1977 (after that date utility was eliminated as a requirement and the requirement of "industrial application" was substituted). *E.g.*, Gold & Shortt, *supra* note 56, at 50. The promise doctrine in Canadian law can arguably be found as early as the 1960s. *E.g.*, *New Process Screw Corp. v. P.L. Robertson Mfg Co. Ltd.* (1961), 39 C.P.R. 31 (Can. Ex. Ct.) (patent promising a process of making screws of many sizes depending on certain pitch angles in fact failed to create the types of screws promised); *see also* *Consolboard Inc. v. MacMillan Bloedel (Sask.) Ltd.*, [1981] 1 S.C.R. 504 (Can.) (stating that an invention is not useful either if it will not operate at all or if "it will not do what the specification promises that it will do"); *see also* Gold & Shortt, *supra* note 56, at 38. *But see* Norman Siebrasse, *The False Doctrine of False Promise*, CANADIAN INTELL. PROP. REV. 3, 8–9 (2012).

88. Siebrasse, *supra* note 87, 36–37. For the purpose of analyzing Eli Lilly's investment claims, this article assumes that Eli Lilly is correct in asserting that the promise doctrine modifies Canadian patent law to consider the legal and policy implications; however, if Canada's laws have not changed at all, Eli Lilly's claims are weaker. Although it is currently unknown how the tribunal will decide, Canada as well as some scholars dispute that Canada's law has changed. *E.g.*, *Eli Lilly & Co. v. Canada*, Statement of Defense, ¶ 12, (NAFTA/UNCITRAL Arb. Trib. June 30, 2014) [hereinafter *Canada, Statement of Defense*]; Gold & Shortt, *supra* note 56, at 53–57, 61–77 (arguing that the promise doctrine is not new to Canada and that other countries also require similar evidence, even if they do so under other patent law requirements).

89. *E.g.*, *Eli Lilly Canada Inc. v. Novopharm Ltd.*, [2012] 1 F.C.R. 349, para. 76 (Can.).

be demonstrated in the patent or “soundly predicted.” In the many cases where the promise relies on a sound prediction, there are three components to satisfy. First, there must be a factual basis for the prediction. Tested compounds can supply this. Second, the inventor must have a sound basis from which the desired result can be inferred from the factual basis as of the date of the application. Third, there must be proper disclosure in the patent application to justify the quid pro quo of a patent monopoly.⁹⁰

An important issue is where a court (or the patent office) should look to find a promise in the patent or in the patent application. Some courts and commentators argue that only promises found entirely or mostly in the claims are relevant, but not those found in the specification.⁹¹ The majority of cases do not restrict promises to those found in patent claims.⁹² This is consistent with claim drafting; patent prosecutors generally do not restrict claims by stating the claims’ utility. Rather, prosecutors draft claims to simply cover the structural elements of the invention. More typically, a patent or patent application will state in its specification the invention’s intended purpose, as well as how the invention is an improvement over prior inventions, thus leading to a promise. Courts have not found promises based solely on abstract tables of data or drawings and instead generally look to an actual statement in the specification.⁹³ However, even focusing on statements in the specification can lead to different results, with some recent decisions since Eli Lilly filed its notice of arbitration seeming to take a more restrained view of what constitutes a promise.⁹⁴

90. *Id.* at para. 70.

91. *E.g.*, *Fournier Pharma Inc. v. Canada (Health)*, 2012 FC 740, para. 126 (Can.); *Teva Canada Ltd. v. Novartis AG*, 2013 FC 141, paras. 76–77 (Can.); Fiona E. Legere, *The Pitfalls of “the Promise of the Patent,”* 29 CANADIAN INTEL. PROP. REV. 57, 60–61 (2013). However, some courts assert that claims control but actually look at the specification. *E.g.*, *Bauer Hockey Corp. v. Eaton Sports Canada Inc.*, 2010 FC 361, para. 289 (Can.).

92. *See* Gold & Shortt, *supra* note 56, at 44. Some Canadian cases decided after Eli Lilly’s notice of arbitration have taken a more restrained view of the promise doctrine. *E.g.*, *Sanofi-Aventis v. Apotex Inc.*, 2013 FCA 186, para. 49 (Can.) (finding that there should only be a promise if “the inventor makes an explicit promise of a specific result” in a case involving the drug sold as Plavix).

93. *E.g.*, *Apotex Inc. v. H Lundbeck A/S*, 2013 FC 192, paras. 244–53 (Can.) (rejecting promise based simply on table of data); *Pfizer Canada Inc. v. Canada (Health)*, [2009] 1 F.C.R. 253, paras. 54–55 (Can. Fed. Ct. App.) (reversing trial court finding of promise based solely on data in a table); *Gold v. Serratus Mountain Prods. Ltd.*, 2004 FC 815, para. 53 (Can.) (drawings alone inadequate).

94. *E.g.*, *Sanofi-Aventis v. Apotex Inc.*, 2013 FCA 186 (Can.) (concluding that the lower court had erred in finding a promise for use in humans based on inferences regarding the drug sold as Plavix); *Pfizer Canada Inc. v. Mylan Pharm. ULC*, 2014 FC 38, para. 60 (Can.) (finding the patent relating to the drug marketed as Celebrex did not promise reduced

Eli Lilly and other companies criticize the promise doctrine as discriminatory since courts have mostly invalidated pharmaceutical patents under it.⁹⁵ However, the doctrine is technically not limited to pharmaceuticals. The Canadian Manual for Patent Practice in fact has a non-pharmaceutical example involving a golf club. In addition, a court has recently applied the doctrine to a mechanical invention.⁹⁶ Even though the promise doctrine applies to all areas, it may be harder to satisfy a promise of an improved medical treatment than an improved mechanical device—especially if the patent promises that the invention is superior for long-term treatment, since proving that claim involves clinical data. However, this does not necessarily support Eli Lilly’s claim of discrimination since a number of neutrally worded patent standards are more difficult to meet for certain areas of technology, yet are not considered discriminatory.

Some also criticize the promise doctrine as being unpredictable. For example, in *Apotex Inc. v. Pfizer Canada Inc.*, the court found the patent application promised to treat glaucoma with minimal side effects; because the court determined that glaucoma was a chronic disease, it invalidated the patent for lacking any long-term data on treatment, even though the patent claims did not refer to a chronic condition or long-term treatment.⁹⁷ This decision also conflicted with an earlier court decision holding the same patent valid.⁹⁸ Although these different decisions may seem to indict the doctrine, courts have come to different conclusions concerning the same

side effects in humans); *Apotex Inc. v. Pfizer Canada Inc.*, 2014 FCA 250, paras. 60–70 (Can.) (rejecting the suggestion that there was a promise of utility in treating inflammation or a promise of reduced side effects in a patent covering the drug marketed as Celebrex in contrast to an earlier decision by a different drug concerning the same patent); *Pfizer Canada Inc. v. Apotex Inc.*, 2014 FC 314 (Can.). *But see* *Alcon Canada Inc. v. Cobalt Pharm. Co.*, 2014 FC 149 (Can.) (finding implied promise of enhanced physical stability of solution). Of course, it could also be that some promises alleged to exist by a company desiring to invalidate the patent were farfetched. For example, in one case, the defendant claimed a passage suggested that a patent promised rapid absorption of the drug in vivo based on language that simply noted that this was a possible mechanism, such that the court rejected this argument. *Bayer Inc. v. Cobalt Pharm. Co.*, 2013 FC 1061 (Can.). Nonetheless, the court stated that “[c]ourts should not strive to defeat otherwise valid patents,” which could be seen as an intent to take a more restrained view of applying the promise doctrine. *Id.* at para. 93.

95. *E.g.*, PHARMA, SPECIAL 301 SUBMISSION 2014 76–77 (2014); *see also* OFFICE OF THE U.S. TRADE REPRESENTATIVE, EXEC. OFFICE OF THE PRESIDENT, 2013 SPECIAL 301 REPORT 46 (2013) (noting “serious concerns about the impact of the heightened utility requirements for patents”).

96. *Bell Helicopter Textron Canada Limitée v. Eurocopter*, 2013 FCA 219 (Can.).

97. 2011 FCA 236 (Can.).

98. *Pharmascience Inc. v. Pfizer Canada Inc.*, 2011 FCA 102 (Can.).

patent in applying other patentability standards. In fact, for some issues, such as claim construction, reversals are quite common.

Although commentators have criticized the promise doctrine for being without basis, the Canadian Supreme Court has provided a firm foundation for the doctrine in public policy. The court explained that it “balances the public interest in early disclosure of new and useful inventions, even before their utility has been verified by tests . . . and the public interest in avoiding . . . granting monopoly rights in exchange for misinformation.”⁹⁹ Noting that patent monopolies are associated with higher prices, the court stated that the “public should not be expected to pay an elevated price in exchange for speculation.”¹⁰⁰

In the case of Eli Lilly’s patents on the drugs sold as Strattera and Zyprexa, courts found the patents on both drugs promised certain treatments, yet failed to soundly predict them, such that the patents were invalid for lack of utility. Moreover, in both cases, Eli Lilly appeared to be attempting to obtain additional patent protection beyond the term of its original patents. In other words, Eli Lilly was engaging in the common practice of “evergreening” by pharmaceutical companies; as noted earlier, some scholars and policymakers suggest that this is precisely the situation where countries should decline to provide patents. The facts of Eli Lilly’s patents further underscore this.

A federal court invalidated Eli Lilly’s improvement patent, which allegedly covered the drug sold as Strattera, because it had an implied promise to treat ADHD as a chronic condition, but the patent only disclosed a short-term study and did not demonstrate efficacy for long-term use.¹⁰¹ The invalidated patent claimed a new use—treating ADHD—for a known compound, and specifically suggested in the specification that the compound was effective without disclosing any studies or working examples. Importantly, Eli Lilly had previously obtained a patent for the broader “genus” (group) of compounds, as well as a second Eli Lilly patent for treatment for depression.¹⁰² Eli Lilly later developed data to establish that the drug was in fact useful for treating ADHD, but not until long after filing the patent application. Admittedly, it is difficult for a patent applicant to have

99. *E.g.*, NAFTA, *supra* note 36, art. 1100; *see also* Apotex Inc. v. Wellcome Found., Ltd. [2002] 4 S.C.R. 153, para. 66 (Can.).

100. *Id.* at para. 37.

101. Novopharm Ltd. v. Eli Lilly & Co., 2010 FC 915 (Can.).

102. Canada Patent No. 1,051,034 (Filing date Jan 1, 1975); Canada Patent No. 1,181,430 (Filing date Nov. 12, 1981); *see also* Canada, Statement of Defense, *supra* note 88, ¶ 53.

clinical data at the time an invention is disclosed. However, the invention at issue here did not involve a new compound; rather, it was a new use of an old compound. Moreover, as noted earlier, patent policy dictates that the public should not be burdened with the social cost of a patent unless, at the time of filing, the inventor has made an adequate disclosure; otherwise, the patent becomes a mere hunting license, which imposes substantial costs for the public while rewarding applicants for making a lucky guess.

Similarly, a different federal court invalidated Eli Lilly's patent claiming use of the drug sold as Zyprexa to treat schizophrenia because the patent implied a promise of superiority. In that case, the court found the patent promised the drug would have fewer side effects than existing antipsychotics for long-term treatment, but the specification provided inadequate disclosure to support this promise.¹⁰³ Importantly, the patent at issue was a "selection" patent that Eli Lilly sought after obtaining a patent on a broader "genus" patent for use in the treatment of certain psychotic conditions, including schizophrenia.¹⁰⁴ However, Canadian patent law, consistent with other countries, generally only permits patents on selections of a previously patented genus if the narrower claim has an advantage over the previously disclosed genus. Eli Lilly's invalidated patent attempted to demonstrate such an advantage; it in fact stated that the invention had "marked superiority and a better side effects profile than prior known antipsychotic agents."¹⁰⁵ However, it did not disclose data to support this claim.

Eli Lilly asserts that Canada improperly invalidated its patents based on an interpretation of the law that did not exist when the patents were examined.¹⁰⁶ Eli Lilly notes that Canada's law is currently different than that of other NAFTA parties (the United States and Mexico) but that when NAFTA was enacted, Canadian law was more similar to other NAFTA parties, such that the promise doctrine could not have been anticipated.¹⁰⁷ Accordingly, Eli Lilly asserts that Canada was wrong to "re-interpret a core patentability requirement enshrined in NAFTA in a way that contradicts the standard accepted by the NAFTA parties at the time the treaty was negotiated."¹⁰⁸

103. *Eli Lilly Canada Inc. v. Novopharm Ltd.*, 2011 FC 1288, para. 218 (Can.).

104. *See* Canadian Patent No. 1,075,687. A selection patent is a subset of the previously and more broadly disclosed genus.

105. Canada Patent 2,041,113, p. 5; *see also* Canada, Statement of Defense, *supra* note 88, ¶ 69 (citing patent specification).

106. *Eli Lilly Notice of Arbitration*, *supra* note 2, ¶ 69.

107. *Id.* ¶¶ 28–34.

108. *Id.* ¶ 68.

Eli Lilly also suggests that the promise doctrine is inconsistent with the PCT because the PCT prohibits countries from “imposing ‘requirements as to the form or contents . . .’” of the original PCT application.¹⁰⁹ In particular, it asserts that the promise doctrine “would defeat the single application objective.”¹¹⁰ Eli Lilly seems to make two separate, but both invalid, claims about the PCT. First, Eli Lilly suggests that the promise doctrine is inconsistent with the PCT requirement that bars countries from imposing requirements on the form of the original PCT application.¹¹¹ Although there is such a rule, it notably only governs the actual PCT application and not patent standards of individual countries. The PCT explicitly states that it does not “limit the freedom of each Contracting State to prescribe such substantive conditions of patentability as it desires.” Utility is in fact a substantive condition of patentability, and the PCT states that national laws may require the applicant to furnish evidence of any substantive condition of patentability.¹¹²

Eli Lilly’s claim that the promise doctrine “would defeat the single application objective” is also flawed. Eli Lilly seems to presume that if one nation has a patent standard that differs from others it would be impossible for applicants to comply with this standard in the PCT application, such that this inconsistent standard would defeat the purpose of the PCT—to provide a streamlined method of applying for patents in multiple countries. However, given that the PCT does not govern patentability standards, this claim is unfounded. This argument seems especially unjustified since there are prior situations where nations have had different laws that implicate what is disclosed in the PCT application. For example, the United States is unique in requiring patent applicants to disclose the “best mode” of using an invention.¹¹³ Although other countries have not adopted the best mode requirement, foreign applicants have used PCT applications to seek U.S.

109. *Id.* ¶ 45; *see also* Patent Cooperation Treaty, *supra* note 51, art. 27(1).

110. Eli Lilly Notice of Arbitration, *supra* note 2, ¶ 45.

111. Patent Cooperation Treaty, *supra* note 64, art. 27(5).

112. *Id.* art. 27(5)–(6).

113. 35 U.S.C. § 112 (2012); *see also* NAT’L RESEARCH COUNCIL OF THE NAT’L ACADEMIES, A PATENT SYSTEM FOR THE 21ST CENTURY 121 (Stephen A. Merrill et al. eds., 2004) (“[O]nly the United States imposes a best mode requirement.”); Donald S. Chisum, *Best Mode Concealment and Inequitable Conduct in Patent Procurement: A Nutshell, A Review of Recent Federal Circuit Cases and a Plea for Modest Reform*, 13 SANTA CLARA COMPUTER & HIGH TECH. L.J. 277, 279 (1997) (noting the standard is “virtually unique to U.S. law” without counterparts in European or Japanese law). However, although this requirement still technically exists for granting of patents, failure to meet this requirement is no longer a basis for invalidation after the most recent amendments to patent laws. 35 U.S.C. § 282(b)(3)(A) (2012).

patents for years. Admittedly, some suggest that this may be unfair to foreign applicants who might forget to consider U.S. patent laws when they file a PCT application. Yet even those that criticize the best mode as unfair to foreign applicants do not cite any problems with PCT applications.¹¹⁴ This criticism differs from Eli Lilly's claim that the promise doctrine is an actual violation of the PCT or that the doctrine completely defeats the single application objective. For example, the best mode requirement does not defeat the single application objective because an applicant's PCT application can still result in a U.S. patent if the applicant discloses the information appropriate for complying with U.S. law.

Although Eli Lilly correctly considers a patent to be a possible investment covered by investor-state agreements, there are multiple challenges to its claims that this Article will detail in subsequent Parts. First, although valid intellectual property rights are unquestionably investments, an intellectual property "right" that is canceled for failing to meet the applicable standards should not be considered an investment that is within the scope of such agreements. Eli Lilly obviously contests this issue in making its claims that invalidating a patent is tantamount to "expropriation" of its investment and a violation of the requirement that all investments be provided "fair and equitable treatment."

II. REVOCATION OF INTELLECTUAL PROPERTY RIGHTS SHOULD NOT BE A BREACH OF INVESTMENT OBLIGATIONS

Eli Lilly's case illustrates why revocation of patent rights should not constitute a breach of investment obligations. First, this Part explains why there is no covered "investment" and thus no grounds for an investor-state dispute. Then, this Part argues that tribunals should not find revoked rights to be either an expropriation or a violation of fair and equitable treatment. Although there are strong reasons for rejecting such claims, this analysis also highlights how a tribunal could nonetheless find otherwise, which underscores the need for the reforms proposed in Part IV.

114. *E.g.*, NAT'L RESEARCH COUNCIL, *supra* note 113, at 121; Chisum, *supra* note 113, at 279. In particular, these criticisms focus on the situation where an applicant relies on the PCT application, but reliance on a previously filed domestic application in the foreign inventor's home country for obtaining an earlier filing date. *See* 35 U.S.C. § 119 (2012). However, those who complain about the unfairness of the best mode for foreign applicants mostly focused not on PCT applications, but applicants who rely on the date of a patent application in their *own country*. *E.g.*, Lee Petherbridge & Jason A. Rantanen, *The Pseudo-Elimination of Best Mode: Worst Possible Choice?*, 59 UCLA L. REV. DISCOURSE 170, 171 (2012).

A. CANCELED INTELLECTUAL PROPERTY RIGHTS SHOULD NOT BE A COVERED INVESTMENT THAT WOULD ENABLE AN INVESTOR-STATE DISPUTE

Existing investment agreements should be interpreted to exclude canceled intellectual property rights as a covered investment. Intellectual property rights are different than other types of property because they can be and often are later canceled. The cancellation of the rights means there were no legitimate rights to begin with, so in these cases there should be no recognized investment that would trigger the ability to file an investor-state dispute.

1. *Intellectual Property is Different from Real Property*

Intellectual property rights are fundamentally different from real property rights with respect to their existence; intellectual property may be canceled and has a different creation process.¹¹⁵ This is inherently different from real property, which is never canceled; in the rare case where an action to quiet title succeeds, the property itself still exists. In addition, unlike most forms of real property, which exist without state intervention, some types of intellectual property only exist if granted by the state and states can even cancel many types that exist without state intervention.¹¹⁶ For example, a patent right does not exist without a state agency such as the U.S. Patent and Trademark Office reviewing an application to evaluate whether a patent is deserved.¹¹⁷ But this determination occurs after only a brief administrative review, so patents are at most presumptively valid;¹¹⁸ they can be and often

115. Although not all intellectual property rights may be canceled, this is definitely true of patents, as well as trademarks and copyrights. *E.g.*, 15 U.S.C. § 1064 (2012) (trademarks); 37 C.F.R. § 201.7 (2001) (copyright).

116. The only types of intellectual property rights that cannot be canceled are rights of publicity and trade secrets. However, these are often considered less valuable to companies than the traditional types of intellectual property that can be canceled, such as patents, copyrights, and trademarks.

117. *E.g.*, 35 U.S.C. §§ 111, 131 (2012). This is obviously different than other types of intellectual property rights, such as trademarks, trade secrets, and rights of publicity, which not only exist, but which a rights holder can enforce without a state determination to grant. *E.g.*, 15 U.S.C. § 1125 (2012) (permitting enforcement of unregistered trademarks). Copyrights can also exist without state determination, but U.S. law still requires registration of a copyright before a rights holder can file suit. 17 U.S.C. § 411 (2012).

118. *E.g.*, 35 U.S.C. § 282 (2012) (presumption of validity); Canada Patent Act, R.S.C. 1985, c. P-4, § 43(2). In India, patents do not enjoy an initial presumption of validity; they only acquire such a presumption after six years, after the opposition period has elapsed. *E.g.*, The Patents Act, No. 39 of 1970, INDIA CODE § 13.4 (2005) (as amended by The Patents Act 2005, No. 15, Acts of Parliament, 2005) (patent examination does not “warrant the validity of any patent”); 3M Innovative Props. Co. v. M/S Venus Safety & Health Pvt Ltd.,

are canceled or revoked if it is found that the agency should not have issued the rights in the first instance.¹¹⁹ In most countries, there are a variety of cancellation and revocation mechanisms, including proceedings at the patent office as well as invalidation in a court.¹²⁰ For example, Canada's patent laws state that an issued patent is assumed valid in the absence of evidence to the contrary, thus expressly contemplating that issued patents can be found invalid.¹²¹ Canada's patent laws also provide that the usual patent right to exclude is "subject to adjudication" by Canadian courts, which means that those rights are contingent on a Canadian court determining whether the patent is valid.¹²² Accordingly, a patent that is invalidated for failure to satisfy one of the stated standards should not constitute intangible property pursuant to an investment agreement since the invalidation means the patent never should have existed.

Rather, invalidating a patent is more akin to an application for patent rights, for which there has never been a recognized property right.¹²³ In particular, whereas the owner of a patent can exclude others from use of the patented invention, the owner of a mere patent application has no such rights.¹²⁴ Nonetheless, there is one prior case that is somewhat analogous to

(2014) (India), available at <http://lobis.nic.in/dhc/MAN/judgement/05-06-2014/MAN30052014S25582013.pdf>; see also TABREZ AHMAD & INDRANIL GHOSH, PRESUMPTION OF PATENT VALIDITY: A COMPARATIVE STUDY IN THE LIGHT OF INDIAN AND U.S. SITUATIONS (2011), available at <http://ssrn.com/abstract=1978918>; Lakshmikumaran & Sridharan, *International Report- Delhi High Court Rules on Presumption of Validity Under Indian Patent Law*, IAM (July 2, 2014) <http://www.iam-magazine.com/reports/Detail.aspx?g=2805840f-8901-40ac-b907-e2ef68e69ceb/>.

119. See, e.g., 35 U.S.C. §§ 311, 321 (2012); Canada Patent Act, R.S.C. 1985, c. P-4, §§ 42, 43(2), 60(1).

120. In the United States, patents may be invalidated pursuant to re-examination, inter partes review, and a relatively new post-grant review for patents filed after March 2013. 35 U.S.C. §§ 301 (re-examination), 311 (inter partes review), 321 (post-grant review) (2012).

121. Canada Patent Act, R.S.C. 1985, c P-4, § 43(2).

122. *Id.* § 42.

123. Although some agreements consider patent applications to be investments, they notably limit such claims to applications for *patentable* inventions, which means that they still must meet the basic patentability requirements. *E.g.*, Bilateral Investment Treaty, U.S.-Jam., art. I.1(a)(iv), Feb. 4, 1994, S. TREATY DOC. NO. 103-35 (1997). However, including an application as an investment seems questionable based on intellectual property laws because there are no rights unless and until they are granted. *But see* Mercurio, *supra* note 26, at 878-80 (arguing that an application for an intellectual property right could be considered an investment, in part because the European Court of Human Rights held that an application for a trademark is a property right for purposes of the European Convention on Human Rights).

124. See 35 U.S.C. § 271(a) (2012) (only providing enforcement rights to patent owner). Although Congress recently amended the Patent Act to provide limited retroactive damages

the situation of cancelled intellectual property. A recent decision held that an application to sell a generic drug was not a property right that would qualify as an investment because, even though the FDA granted tentative approval, the FDA retained the right to revoke the approval.¹²⁵ Although an application to sell a drug is obviously different than an application for a patent, in both situations there is no “right” (whether to sell a drug or exclude others from making it) unless and until a government agency makes a determination. In addition, in both cases, a positive government decision is tentative and subject to reversal. Since a revocable decision that would confer some proper rights is not an investment, a right that has already been revoked is even less likely to qualify. Of course, this is only one tribunal decision, and other decisions need not follow it. However, given that the decision simply further supports long-standing patent policy, there is a firm foundation for a tribunal to consider cancelled intellectual property rights not to be a type of covered investment.

In addition, canceled intellectual property rights, including those canceled based on a common law modification of long-standing patent criteria should be considered to never have existed. This highlights yet another difference between intellectual property law and real property. As noted earlier, real property is not dependent on state determinations, so what constitutes valid real property is unlikely to change. This difference is an important nuance for Eli Lilly’s case. Eli Lilly argues that because its patents were consistent with Canadian law at the time of application, it was improper for a change in the law after issuance to invalidate its patents.¹²⁶ However, contrary to Eli Lilly’s assertion, it is common for case law to modify patent law and retroactively invalidate previously granted patent rights. In the United States, for example, the Supreme Court’s modification of the obviousness standard to make it

to a patent owner for infringement that occurred while the application was pending, such rights only issue if a patent is ultimately awarded. *See id.* § 154(d).

125. *Apotex Inc. vs. United States*, Award on Jurisdiction and Admissibility, ¶¶ 209–10, 222–23 (NAFTA/UNCITRAL Arb. Trib. June 14, 2013).

126. *Eli Lilly Notice of Arbitration*, *supra* note 2, ¶¶ 8–9 (asserting that the Canadian judiciary has “created a new doctrine” to assess utility that is a “dramatic departure from the standard” prevailing in Canada when its patents were filed and granted). Eli Lilly bases this statement on the fact that when it applied, the Canadian patent office guidelines used a different utility standard that considered inventions to satisfy unless the invention is “totally useless.” *Id.* ¶ 8. However, the very guidelines that Eli Lilly uses as its sole source of authority for this issue in fact state that only courts have authority to interpret patent law. *Manual of Patent Office Practice*, Consumer and Corporate Affairs Canada, Patent Office, Forward (1977). This is true of all subsequent guidelines. *Canada, Statement of Defense*, *supra* note 88, ¶ 46.

more difficult to meet impacted the validity of existing patents.¹²⁷ More recently, after the Supreme Court modified what types of genes may be patentable subject matter, the validity of some previously issued patents is in question.¹²⁸ Moreover, this phenomenon is common to all areas of common law doctrine that have both prospective and retrospective application.¹²⁹ In addition, there is even a prior tribunal decision rejecting the suggestion that retroactive application of a domestic law would be inappropriate and explicitly noting that it is “normally a matter of local courts to determine” whether to apply new decisional law retrospectively.¹³⁰

2. *Canceled Intellectual Property Should Be Excluded from Investment Agreements Based on Policy Grounds*

In addition to the criticisms of investor-state arbitration disputes previously noted, these disputes present unique policy problems when investors challenge IP rights that are permissible under international agreements. First, unlike most types of investments subject to investor-state disputes, international treaties, such as TRIPS, govern intellectual property. These international agreements represent negotiated norms among states, and a decision by an investor-state tribunal interpreting these treaties, or even simply the filing of an action challenging them, would have a chilling effect on these negotiated norms and could also result in inconsistent decisions. Second, intellectual property rights are granted to effectuate domestic social policies, and permitting challenges to domestic decisions canceling intellectual property rights undermines these policies. Third, intellectual property is also fundamentally different from traditionally protected investments in that the traditional rationale for permitting investors to challenge states does not apply to intellectual property.

127. *KSR Int'l Co. v. Teleflex, Inc.*, 550 U.S. 398, 407 (2007); *see also* Christopher M. Holman, *Unpredictability in Patent Law and Its Effect on Pharmaceutical Innovation*, 76 MO. L. REV. 645, 681–84 (2011).

128. *E.g.*, *Ariosa Diagnostics, Inc. v. Sequenom, Inc.*, 19 F. Supp. 3d 938 (N.D. Cal. 2013); *see also* Sherman Kahn, *Will Patents Be the Next Wave in Investor-State Arbitration?*, 7 N.Y. DISP. RESOL. LAW. 53, 56 (2014).

129. *See, e.g.*, Harold J. Krent, *The Puzzling Boundary Between Criminal and Civil Retroactive Lawmaking*, 84 GEO. L.J. 2143, 2156 (1996); *see also* Donald T. Hornstein, *Resiliency, Adaptation, and the Upsides of Ex Post Lawmaking*, 89 N.C. L. REV. 1549, 1551 (2011) (noting that retroactivity is not only tolerated but sometimes celebrated).

130. *Mondev Int'l Ltd. v. United States*, ICSID Case No. ARB(AF)/99/2, Award, ¶¶ 137–38 (Oct. 11, 2002) (rejecting the suggestion that retrospective application of a rule by a court would constitute violation of NAFTA article 1105 regarding minimum standard of treatment).

- a) International Agreements Permit Nations to Decide on the Scope of Intellectual Property Rights; Investment Arbitrations Should Not Disrupt This Norm

Investment arbitration disputes that affect domestic decisions concerning intellectual property rights are different than most other investment disputes because they can undermine the goals of the separate international agreement, TRIPS; accordingly, investor-state disputes should exclude domestic intellectual property decisions consistent with TRIPS.¹³¹ As noted earlier, well over one hundred countries, including Canada, must comply with TRIPS. However, TRIPS sets minimum, but not uniform standards, such that nations have substantial flexibility to define these standards.¹³² In particular, while patents must be granted on all inventions that satisfy traditional criteria, including that the invention be useful, TRIPS notably does not define the term, so countries have discretion to decide the applicable standard.¹³³ In addition, TRIPS expressly contemplates that patent rights can be revoked and simply requires that there be judicial review of any such decision.¹³⁴ As explained below, TRIPS is fundamentally different from other international agreements implicated in investor-state disputes to date, which involve more ambiguous criteria and do not create a direct conflict. In addition, if investor-state disputes could challenge TRIPS-consistent decisions, there is a risk of decisions inconsistent with the built-in dispute resolution process of TRIPS. Moreover, considering that TRIPS is already an encroachment on traditional state sovereignty, permitting investor-state disputes to challenge TRIPS-consistent actions would seem particularly unfair and would have a chilling effect on TRIPS-permissible laws that would promote better access to affordable medicine. Given these issues, investor-state tribunals should not consider TRIPS-consistent domestic decisions as violations of investor rights.

Although investor-state tribunals have previously addressed conflicts with other international agreements, those agreements are fundamentally different than TRIPS. Only in limited situations has another international agreement been arguably relevant to investment-specific claims.¹³⁵ These

131. Although there are agreements since TRIPS that set even higher standards, this section will focus only on TRIPS because it has the most extensive membership of any international agreement concerning intellectual property.

132. TRIPS, *supra* note 3, art. 1(1).

133. *Id.* art. 27(1); *see also supra* notes 53–60 and accompanying text (noting that this proposition is well supported by scholars and policymakers).

134. TRIPS, *supra* note 3, art. 32.

135. This excludes the use of “umbrella” clauses in international agreements that permit an investor to enforce other commitments or reliance on other international agreements as

situations involve direct conflicts with another international agreement where arguably only one agreement can apply,¹³⁶ and a situation where an agreement that generally provides for investor-state arbitration explicitly exempts under certain environmental treaties.¹³⁷ TRIPS, however, is a unique international agreement because its standards allow for flexibility. An investor-state tribunal could find that a state has violated the investment-based rights of a foreign investor when the state is in compliance with TRIPS because TRIPS only provides a minimum standard of protection. Although this may seem to suggest there is no conflict, such a finding would undermine the ability of countries to use flexibilities under TRIPS. As a policy matter, it is inappropriate to impose liability under an investment chapter on a country that is complying with a separate international agreement.

In no other prior situations have investor-state disputes had the potential to create liability for nations that are complying with a separate international agreement. Indeed, in most cases, there is not a true conflict with a separate international agreement. The closest situation that has arisen involves the

part of customary international law. *See, e.g.*, Roger Alford, *Using Investment Arbitration to Enforce WTO Commitments*, KLUWER ARB. BLOG (Apr. 18, 2014), <http://kluwarbitrationblog.com/blog/2014/04/18/using-investment-arbitration-to-enforce-wto-commitments/> (noting that umbrella clauses in BITS could permit enforcement of WTO and other trade agreement commitments); *see also* Charles Owen Verrill, Jr., *Are WTO Violations Also Contrary to the Fair and Equitable Treatment Obligations in Investor Protection Agreements?*, 11 ILSA J. INT'L & COMP. L. 287 (2005); *S.D. Myers, Inc. v. Canada*, Partial Award, ¶¶ 234, 256 (NAFTA/UNCITRAL Arb. Trib. Nov. 12, 2000), 40 I.L.M. 1408 (2001) [hereinafter *S.D. Myers Partial Award*] (suggesting in dicta that minimum standard of treatment under NAFTA take into account “the letter or spirit of widely, though not universally accepted international agreements like those in the WTO system”).

136. Even when raised, tribunals do not always need to address a possible conflict since sometimes the competing international agreement is found to have come into force too late to be relevant. *E.g.*, *Micula v. Romania*, ICSID Case No. ARB/05/20, Award (Dec. 11, 2013) (finding no conflict because non-investment agreement was concluded *after* the investment agreement). *See also* Abba Kolo, *Transfer of Funds: The Interaction between the IMF Articles of Agreement and Modern Investment Treaties: A Comparative Law Perspective*, in INTERNATIONAL INVESTMENT LAW AND COMPARATIVE PUBLIC LAW 345, 355 (Steven Schill ed., 2010) (suggesting that if there is a conflict between a capital transfer provision of an investment agreement and the IMF, the investment agreement should apply as more specific under *lex specialis*).

137. Canada argued that investment claims were inconsistent with its international environmental obligations under the Basel Convention and the U.S.-Canada Transboundary Agreement on Hazardous Waste, such that these agreements should prevail over investment claims. The tribunal found no actual conflict. *S.D. Myers Partial Award*, *supra* note 135, ¶ 150. This situation is also different in that NAFTA has a specific clause for possible inconsistency with other environmental agreements and states that those obligations shall prevail to the extent of inconsistency, but where a party has a choice between equally effective and reasonably available means, the party must choose the least inconsistent measure. NAFTA, *supra* note 36, art. 104.

actual or proposed use of other international agreements as defenses where those agreements are largely aspirational. For example, in the pending suits concerning plain packaging tobacco laws, there is a WHO framework convention that supports domestic laws at issue,¹³⁸ but legally, no member countries must apply the guidelines.¹³⁹ Similarly, although some commentators have suggested that states should rely on international human rights norms as a defense, these norms are notably vague and generally do not clearly indicate what countries can do.¹⁴⁰ For example, although some have suggested that there is an international right to affordable and equitably distributed water that could be relevant, no state has attempted to clearly rely on such rights.¹⁴¹ This could be because international rights are to citizens, rather than countries. Moreover, there is no firm requirement; any such right is based on a General Comment to the UN Committee on Economic, Social and Cultural Rights, and the UN agreement actually only requires that states “take steps” with the caveat that those steps be based on “available resources.”¹⁴² Accordingly, as human rights treaties simply reflect a general desire to promote certain activity, they provide no firm criteria, whereas TRIPS expressly states that countries must provide certain patent standards while at the same time permitting nations flexibility in interpreting these standards.

138. *E.g.*, Philip Morris Brands Sàrl v. Oriental Republic of Uru., ICSID Case No. ARB/10/7, Request for Arbitration, ¶ 3 (Feb. 19, 2010); Philip Morris Asia Ltd. v. Australia, PCA Case. No. 2012-12, Written Notification of Claim (BIT/UNCITRAL Arb. Trib. June 27, 2011).

139. *E.g.*, Susy Frankel & Daniel Gervais, *Plain Packaging and the Interpretation of the TRIPS Agreement*, 46 VAND. J. TRANSNAT'L L. 1149, 1163 (2013); *see also* WORLD HEALTH ORG., WHO FRAMEWORK CONVENTION ON TOBACCO CONTROL (2004), *available at* <http://whqlibdoc.who.int/publications/2003/9241591013.pdf>.

140. *E.g.*, LUKE ERIC PETERSON & KEVIN R. GRAY, INT'L INST. FOR SUSTAINABLE DEV., INTERNATIONAL HUMAN RIGHTS IN BILATERAL INVESTMENT TREATIES AND IN INVESTMENT TREATY ARBITRATION 27–32 (2003) (suggesting that states that are member to ICESCR and other covenants consider using these as defenses); Ursula Kriebaum, *Privatizing Human Rights: The Interface Between International Investment Protection and Human Rights*, in THE LAW OF INTERNATIONAL RELATIONS 165, 173–77 (A. Reinisch & U. Kriebaum eds., 2007); Todd Weiler, *Balancing Human Rights and Investor Protection: A New Approach for a Different Legal Order*, 27 B.C. INT'L & COMP. L. REV. 429 (2004).

141. Indeed, one commentator has suggested that the vagueness of these obligations is an explanation for why nations do not suggest a conflict with human rights. James D. Fry, *International Human Rights Law in Investment Arbitration: Evidence of International Law's Unity*, 18 DUKE J. COMP. & INT'L L. 77, 93–96 (2007).

142. International Covenant on Economic, Social and Cultural Rights art. 2, Dec. 16, 1966, 993 U.N.T.S. 3 (“[E]ach State Party to the present Covenant undertakes to take steps, . . . with a view to achieving progressively the full realization of the rights recognized.”); *see also* arts. 11–13 (right to water). In addition, the General Comment is itself nonbinding.

Investor-state dispute tribunals should also decline to address intellectual property issues that are consistent with international agreements concerning intellectual property when these agreements already provide a mechanism for addressing alleged inconsistencies to prevent inconsistent decisions. For example, there is a built-in forum for adjudicating alleged TRIPS violations pursuant to the robust WTO dispute settlement process. If investors were permitted to usurp this process, it could both result in inconsistent decisions and undermine the negotiated international norms pursuant to TRIPS. Notably, the WTO dispute settlement process is intended to be the sole means to settle violations of its agreements such as TRIPS. Although there is no language expressly excluding investor-state arbitrations, no such arbitrations involved intellectual property at the time the WTO and TRIPS were negotiated, so negotiators likely did not see the need to include such a provision. However, these agreements do contain language prohibiting countries from taking unilateral action for violations.¹⁴³ Permitting investors to engage in a form of self-help through investor-state arbitrations seems one step beyond countries taking unilateral actions. Moreover, there are issues with having investor-state arbitrations decide TRIPS issues when they lack familiarity with either intellectual property or WTO agreements.¹⁴⁴ There is a strong possibility of inconsistent rulings, especially because investor-state arbitrations have no appellate review.

Beyond interfering with an existing dispute resolution process and producing potentially inconsistent decisions, permitting investor-state arbitrations to overrule internationally agreed upon domestic flexibilities under TRIPS seems particularly unfair to countries since TRIPS already encroaches on traditional domestic authority in the area of intellectual property rights. Notably, the TRIPS requirement that all countries provide some level of patent protection was a monumental change to the prior international landscape, where countries previously did not have to grant any intellectual property rights. The idea of global rules requiring patent protection was the brainchild of multinational pharmaceutical companies. They successfully lobbied the United States and EU member states to advocate this in the context of an agreement that would include issues of interest to developing countries that would otherwise oppose an agreement

143. Understanding on Rules and Procedures Governing the Settlement of Disputes art. 23:2 Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, 33 I.L.M. 1226, 1241.

144. Indeed, some suggest that past tribunals have struggled to properly interpret and apply WTO law. *E.g.*, Jürgen Kurtz, *The Use and Abuse of WTO Law in Investor-State Arbitration: Competition and its Discontents*, 20 EUR. J. INT'L L. 749 (2009) (focusing on misuse of WTO law concerning national treatment); Mercurio, *supra* note 26, at 905.

focused exclusively on mandating intellectual property rights.¹⁴⁵ Developing countries may have capitulated to including intellectual property norms because they were interested in enhancing the ability to market agricultural products to other countries, which membership in the WTO would allow.¹⁴⁶ In addition, some developing countries may have agreed to TRIPS assuming that this would forestall unilateral pressure from countries concerning their intellectual property laws.¹⁴⁷ Accordingly, agreement to TRIPS requirements, including providing patents on drugs, does not reflect uniform agreement that patents are desirable as a matter of policy. Given this historical context, permitting an individual investor to further encroach on the limited domestic flexibilities under TRIPS seems particularly unfair to developing countries.

Importantly, if tribunals allow cases such as *Eli Lilly's*, these cases could have a chilling effect on an important trend where developing countries are beginning to finally use their full flexibility under TRIPS. Notably, although TRIPS has always provided states discretion to define the minimum patentability standards, some nations were initially hesitant to do so and simply copied the patent standards of countries such as the United States, even though such laws were not necessarily in their interest. India was the first country to use its full flexibility under TRIPS to create a unique law that bars patents on “new” drugs that are in fact only modest variations of old drugs with no improved benefit to patients.¹⁴⁸ Since India adopted its law in 2005, other countries have either copied India’s laws, or are contemplating doing so.¹⁴⁹ For example, there is a proposal for Brazil to amend its patent standards to mirror India’s.¹⁵⁰

145. *E.g.*, SUSAN K. SELL, *PRIVATE POWER, PUBLIC LAW: THE GLOBALIZATION OF INTELLECTUAL PROPERTY RIGHTS* (2003).

146. *E.g.*, PETER DRAHOS & JOHN BRAITHWAITE, *INFORMATION FEUDALISM: WHO OWNS THE KNOWLEDGE ECONOMY?* 10 (2002).

147. *E.g.*, Peter K. Yu, *TRIPS and its Discontents*, 10 MARQ. INTELL. PROP. L. REV. 369, 372–73 (2006).

148. The Patents (Amendment) Act, 2005, No. 15, § 3(d), Acts of Parliament, 2005 (India).

149. *E.g.*, Universally Accessible Cheaper and Quality Medicines Act of 2008, Rep. Act No. 9502, §§ 1–2 (2008) (Phil.); Divya Rajagopal, *EU, Australia, Canada May Follow India's Patent Law*, ECON. TIMES, Apr. 4, 2013. In addition, Argentina’s patent guidelines impose an even higher standard—new forms are not patentable even if they have an increase in efficacy. Joint Resolution Nos. 118/20012, 546/2012 and 107/2012, May 2, 2012, B.O. 08/05/12 (Arg.), available at http://www.wipo.int/wipolex/en/text.jsp?file_id=288443; Park et al., *supra* note 10, at 43 n.7.

150. BRAZIL CENTER FOR STRATEGIC STUDIES AND DEBATES, *BRAZIL'S PATENT REFORM: INNOVATION TOWARDS NATIONAL COMPETITIVENESS* 114 (2013), available at http://infojustice.org/wp-content/uploads/2013/09/Brazilian_Patent_Reform.pdf. In addition, there is a policy proposal recommending that South Africa either copy India’s law

In light of these negative policy implications, tribunals that end up evaluating cases such as Eli Lilly's should at a minimum take TRIPS—including the inherent domestic flexibilities in its implementation—into account. A recent arbitral decision suggests that a tribunal might be willing to do so. In *Ioan Micula v. Romania*, although the majority of the tribunal rejected the notion that the investment agreement was in direct conflict with the EU Agreement, it was willing to consider EU law in evaluating the investment claims.¹⁵¹ In addition, even if a tribunal were to consider TRIPS, the outcome is unclear because TRIPS does not necessarily result in a conflict with any investment provision. Any possible “conflict” would be more with respect to whether an investment tribunal interpreted TRIPS provisions differently than a WTO panel or scholar would. Different interpretations are possible since companies and even countries have been known to suggest TRIPS violations where there are none.¹⁵² Indeed, Eli Lilly's claims are one example—Eli Lilly alleges that Canada's promise doctrine is inconsistent with the utility standard because of a supposed “shared understanding” between signatories.¹⁵³ However, this is irrelevant to interpreting international treaties; the final text, rather than presumed unstated understandings, controls. Eli Lilly ignores this fact, as well as the fact that utility is undefined in the pertinent treaty, rendering Canada free to define it.

b) Domestic Policy Underlying Intellectual Property Supports
Deferring to States

Intellectual property rights are inherently different than most other types of investments protected by investment chapters. The underlying policy goals that justify providing intellectual property rights are distinct from the goals behind other types of property. For example, patents are the primary policy tool to promote innovation and encourage sharing of inventions, rather than keeping them secret. However, it is well recognized that desired policy goals must be balanced against other competing social goals, such as access to

or an even more stringent approach adopted by Argentina. CHAN PARK ET AL., *supra* note 10, at 46.

151. *Micula v. Romania*, ICSID Case No. ARB/05/20, Final Award (Dec. 11, 2013).

152. *E.g.*, *infra* Subsection IV.A.1 (discussing improper suggestions that India's patent law fails to comply with TRIPS).

153. Eli Lilly Notice of Arbitration, *supra* note 2, ¶ 40; *see also id.* ¶¶ 7–9 (arguing that Canada's changed interpretation of utility is inconsistent with NAFTA's utility requirement without acknowledging that the term is nowhere defined in NAFTA). Technically Eli Lilly claims that Canada's patent standards are inconsistent with NAFTA, rather than TRIPS, but the standards are identical. *Compare* TRIPS, *supra* note 3, art. 27(1), *with* NAFTA, *supra* note 36, art. 1. Indeed, Eli Lilly does not dispute this. Eli Lilly Notice of Arbitration, *supra* note 2, ¶ 42.

affordable medicine. Most other property rights do not inherently compromise these other social goals; indeed, traditional property rights generally do not result in higher prices for goods. Accordingly, although TRIPS requires most countries to provide some degree of patent protection for drugs, it explicitly recognizes the importance of considering public health and other policies; in addition, member states subsequently reaffirmed this principle in a WTO Ministerial Declaration referred to as the Doha Public Health Declaration.¹⁵⁴

Although most countries must provide patents on drugs under TRIPS, it is especially important to defer to domestic decisions concerning TRIPS-consistent patent laws now. First, patents on drugs inherently limit short-term access, but nations are no longer at liberty to completely deny such patent protection even if they value access to low cost medicine more than promoting possible future innovation. Moreover, the patent policy of promoting innovation in the drug arena with patent rights should be considered in light of current business realities. Facing a “crisis” in pharmaceutical innovation where innovation has been stagnant despite exponential increases in expenditures on research, drug companies have developed patent and innovation strategies that aim to extend their profits with minimal innovation. For example, companies are patenting slight modifications of existing drugs, such as extended releases, or new uses that are easier to identify than a brand new compound.¹⁵⁵ In addition, companies are also obtaining patents on multiple aspects of a drug including not just the traditionally patented active ingredient, but also the coating of a drug, or the metabolized version in a patient’s stomach. Critics have dubbed both of these practices “evergreening” because the patent term seems perpetual.¹⁵⁶ In addition, policymakers in both developing and well-developed countries have criticized the handling of these patents; for example, Europe issued a substantial report concerning these patents and recent policy reports suggest that developing countries should modify their patent laws to deny these types of questionable innovations.¹⁵⁷ Companies actually recognize that some of

154. TRIPS, *supra* note 3, arts. 7–8; Doha Declaration on Public Health, *supra* note 34, ¶ 4 (2001).

155. *E.g.*, *supra* note 10 and accompanying text.

156. *E.g.*, THOMAS, *supra* note 10.

157. *E.g.*, EC PHARMACEUTICAL SECTOR INQUIRY, *supra* note 10; BRAZIL CTR. FOR STRATEGIC STUDIES & DEBATES, *supra* note 150; AUSTR. GOV’T, PHARMACEUTICAL PATENTS REVIEW: BACKGROUND AND SUGGESTED ISSUES PAPER (2012); *see also* DECLARATION ON PATENT PROTECTION AND REGULATORY SOVEREIGNTY UNDER TRIPS (2014), available at <https://www.mpg.de/8132986/Patent-Declaration.pdf> (signed by forty scholars from over twenty-five countries to reinforce TRIPS flexibilities).

the patents are of dubious validity,¹⁵⁸ but nonetheless seek such patents in hopes of stemming revenue losses as patents on profitable innovative drugs of prior years such as Lipitor and Prozac increasingly expire.¹⁵⁹ Accordingly, although there could theoretically be a policy justification that patent rights promote innovation even if this negatively impacts short-term access through higher prices, that justification seems more theoretical than real in cases where companies are creating minor improvements. Although companies often argue that basic market principles make such concerns irrelevant because consumers would not buy inferior drugs, as noted earlier, the pharmaceutical market is unique, such that general market principles do not apply. So, countries should be able to use their discretion under TRIPS to minimize the social harm of expensive drugs to only those drugs that they deem are more innovative and thus worth the “cost.”

The contested Eli Lilly patents are the very type of patents that policymakers question. In both cases, Eli Lilly is seeking to obtain additional patent protection when it had at least one patent already. In the case of the drug marketed as Strattera for attention deficit disorder, Eli Lilly was already awarded two different patents before it sought the third patent that Canada invalidated.¹⁶⁰ The drug marketed as Zyprexa similarly already enjoyed a full term of patent protection. Both of these instances could be considered examples of evergreening profitable patents. Indeed, Eli Lilly’s two inventions at issue would likely be invalid in India where, to help address this very type of problem, there is a complete bar on patents that simply claim a new utility for a known compound.¹⁶¹ Moreover, other countries including Brazil, Australia, and member states of the European Union similarly

158. *E.g.*, EC PHARMACEUTICAL SECTOR INQUIRY, *supra* note 10, at 192 (noting companies admitting a strategy to seek patents which “might not be rock solid”).

159. Melly Alazraki, *The 10 Biggest-Selling Drugs That are About to Lose their Patent*, DAILY FIN. (Feb. 27, 2011), <http://www.dailyfinance.com/2011/02/27/top-selling-drugs-are-about-to-lose-patent-protection-ready/> (noting that companies can lose up to ninety percent of sales when patents expire). Notably, although Eli Lilly suggests that Canada improperly invalidated a patent pertaining to Zyprexa, which was one of the top five best selling drugs in 2011, in most countries, Zyprexa lost patent protection in 2012. *Id.*; ELI LILLY & CO., 2012 ANN. REP. 12 (2012), *available at* <http://files.shareholder.com/downloads/LLY/3538664350x0x648089/D3A84E25-2AE1-4E41-8E25-A1AE41DA09BB/English.PDF> (noting Zyprexa sales decreased eighty-three percent in the United States and forty-five percent outside the United States due to loss of patent exclusivity).

160. Canada, Statement of Defense, *supra* note 88, ¶ 53.

161. The Patents (Amendment) Act, 2005, No. 15, § 3(d), Acts of Parliament, 2005 (India).

recognize that current patent laws impacting drugs need to be recalibrated to better balance promoting optimal innovation with less social cost.¹⁶²

c) Considering Intellectual Property an Investment Does Not Foster Traditional Policy Justifying Investor-State Disputes

Intellectual property should be excluded from investor-state arbitration because providing enhanced protection of IP does not satisfy traditional justifications for investment arbitrations. Such provisions arose as a means both to encourage investors to consider countries that they might be hesitant to invest in, and to provide a remedy to foreign investors who might otherwise have no recourse if a state took action that reduced the value of their investments.¹⁶³ As explained in this Section, neither of these justifications is relevant to Eli Lilly's case or to intellectual property in general. First, increased protection of intellectual property through investor-state disputes is unlikely to result in greater investment by owners of intellectual property rights in these countries. Second, it is also not necessary to provide adequate recourse. To the contrary, providing such rights would give foreign investors more recourse than domestic investors.

Permitting intellectual property, including denial of intellectual property rights pursuant to domestic law, to be a covered investment is unlikely to encourage companies to invest in particular countries. Multinational companies do not invest in countries solely based on intellectual property laws. A number of studies indicate that other factors, such as tax incentives, infrastructure, and skills are more relevant than intellectual property laws.¹⁶⁴ Indeed, countries known to have weak intellectual property rights, such as India and China, nonetheless have substantial foreign direct investment.¹⁶⁵ In addition, scholars have noted that especially for impoverished countries,

162. *E.g.*, EC PHARMACEUTICAL SECTOR INQUIRY, *supra* note 10; BRAZIL, CENTER FOR STRATEGIC STUDIES & DEBATES, *supra* note 150; AUSTL. GOV'T, *supra* note 157.

163. *E.g.*, Salacuse, *supra* note 64, at 109–10.

164. *E.g.*, Carsten Fink, *Intellectual Property Rights and U.S. and German International Transactions in Manufacturing Industries*, 9 DUKE J. COMP. & INT'L L. 163 (1997); Claudio R. Frischtak, *Harmonization Versus Differentiation in Intellectual Property Right Regimes*, in GLOBAL DIMENSIONS OF INTELLECTUAL PROPERTY RIGHTS IN SCIENCE AND TECHNOLOGY 89, 103–05 (Mitchel B. Wallerstein et al. eds., 1993); Jeong-Yeon Lee & Edwin Mansfield, *Intellectual Property Protection U.S. Foreign Direct Investment*, 78 REV. ECON. & STAT. 181, 181–85 (1996).

165. For example, India had \$1 billion in foreign direct investment in three months of 2013 despite controversial patent laws that have been noted as inadequate by many companies. *E.g.*, *India Receives Highest FDI Worth \$ 1 billion in Pharma in April–June*, ECON. TIMES (Sept. 1, 2013), http://articles.economictimes.indiatimes.com/2013-09-01/news/41663407_1_pharmaceuticals-sector-highest-fdi-fdi-policy/.

foreign direct investment is unlikely because of a lack of infrastructure or a viable domestic market.¹⁶⁶

This is particularly true for pharmaceutical companies and patent rights. Given the extensive infrastructure already present for the development and manufacture of a drug, local laws are unlikely to encourage investment in a new country. Generally, multinational companies develop patentable inventions where they have research labs, primarily in the United States and Europe, but seek patent rights in all nations where they can market their inventions, including nations where they may have made no investments.

Although some claim that stronger patent rights may promote foreign direct investment, there is no robust empirical evidence to support this claim.¹⁶⁷ A number of scholars have noted that macroeconomic factors such as infrastructure and skills are more important than intellectual property protection,¹⁶⁸ and there are specific studies that note this for patents in particular.¹⁶⁹ Historical evidence supports this argument. For example, Brazil and Thailand received substantial foreign direct investment in the 1970s and 1980s despite low levels of intellectual property protection.¹⁷⁰ One scholar found that foreign direct investment even increased in Korea's pharmaceutical industry after abolition of protection of drugs.¹⁷¹ More recently, although South Africa increased patent protection to comply with TRIPS, this reduced foreign direct investment from pharmaceutical

166. KEITH E. MASKUS, INTELLECTUAL PROPERTY RIGHTS IN THE GLOBAL ECONOMY 199–205 (2000).

167. *E.g.*, SISULE F. MUSUNGU, RETHINKING INNOVATION, DEVELOPMENT AND INTELLECTUAL PROPERTY IN THE UN: WIPO AND BEYOND (2005); Mila Kashcheeva, *The Role of Foreign Direct Investment in the Relation Between Intellectual Property Rights and Growth*, 65 OXFORD ECON. PAPERS 699 (2013); Rajnish Kumar Rai, *Effect of TRIPS-Mandated Intellectual Property Rights on Foreign Direct Investment in Developing Countries: A Case Study of the Indian Pharmaceutical Industry*, 11 J. WORLD INTELL. PROP. 404, 411 (2009).

168. *See supra* note 164; *see also* Wolfgang E. Siebeck, *Conclusions and Recommendations*, in STRENGTHENING OF INTELLECTUAL PROPERTY RIGHTS IN DEVELOPING COUNTRIES 89 (Wolfgang E. Siebeck ed., 1990); Marco C.E.J. Bronckers, *The Impact of TRIPS: Intellectual Property Protection in Developing Countries*, 31 COMMON MARKET L. REV. 1245 (1994); Carlos A. Primo Braga, *Trade-Related Intellectual Property Issues: The Uruguay Round Agreement and its Economic Implications*, in URUGUAY ROUND AND THE DEVELOPING ECONOMIES 381 (Will Martin & L. Alan Winters eds., 1996).

169. *E.g.*, EDWIN MANSFIELD, INTELLECTUAL PROPERTY PROTECTION, DIRECT INVESTMENT, AND TECHNOLOGY TRANSFER, INTERNATIONAL FINANCE CORPORATION, THE WORLD BANK, DISCUSSION PAPER NO. 27 (1994).

170. CARLOS CORREA, INTELLECTUAL PROPERTY RIGHTS, THE WTO AND DEVELOPING COUNTRIES: THE TRIPS AGREEMENT AND POLICY OPTIONS 27 (2000).

171. Arman S. Kirim, *Reconsidering Patents and Economic Development: A Case Study of the Turkish Pharmaceutical Industry*, 13 WORLD DEV. 219, 219–36 (1985).

companies that instead consolidated their operations.¹⁷² This is also consistent with a classic study that showed that stronger patent protection in Italy did not result in increased domestic or foreign direct investment.¹⁷³ Also, some have suggested that even if stronger intellectual property rights might attract some foreign direct investment, the corresponding loss in jobs erodes that benefit. Furthermore, stronger intellectual property rights could create higher local prices that may be cost-prohibitive in developing countries.¹⁷⁴

If intellectual property laws will not encourage multinational companies to invest in a country, providing a remedy when the intellectual property laws are considered undesirable does not seem appropriate because the remedy will not result in substantial investments. In addition, it is fundamentally different than the traditional rationale of protecting induced investments. For example, even though Eli Lilly claims that it could not have anticipated that Canada would change its patent laws, those laws did not induce Eli Lilly to develop the inventions on which it sought Canadian patent protection. Rather, Eli Lilly was developing those inventions for any country that would provide protection and does not appear to have made any specific investments in Canada based on Canadian patent law. Admittedly, an issued patent could have induced Eli Lilly to begin to market its drug. However, unlike other investments for which the investor-state dispute challenge is granted, a granted patent can be legally canceled, so the claim for inducement is weak. Moreover, the “investment” of marketing a drug does not provide the same value to a country as the types of investment that the laws initially contemplated. For example, although marketing a drug may involve some investment in local advertisement and possibly employment of local citizens, it would likely pale in comparison to the capital investments that are more typical with real property such as building structures in a country in response to incentive programs.

In addition, investor-state arbitrations originally developed to provide foreign investors an ability to protect assets when they had no other means to do so. Typically, this was because they could not bring a claim before

172. C. te W. Naudé & J.M. Luiz, *An Industry Analysis of Pharmaceutical Production in South Africa*, 44 S. AFR. J. BUS. MGMT. 33 (2013); David Kaplan, *Intellectual Property Rights and Innovation in South Africa: A Framework*, in THE ECONOMICS OF INTELLECTUAL PROPERTY IN SOUTH AFRICA (WORLD INT’L PROP. ORG. ed., 2009).

173. F.M. Scherer & Sandy Weisburst, *Economic Effects of Strengthening Pharmaceutical Patent Protection in Italy*, 26 INT’L REV. INDUS. PROP. & COPYRIGHT L. 1009 (1995).

174. E.g., Paul J. Heald, *Mowing the Playing Field: Addressing Information Distortion and Asymmetry in the TRIPS Game*, 88 MINN L. REV. 249 (2003); Keith E. Maskus, *Intellectual Property Rights and Economic Development*, 32 CASE W. RES. J. INT’L L. 471, 480–81 (2000).

domestic courts where the government might be immune from suit or because court systems were corrupt. However, neither of these situations apply to Eli Lilly's case. It was already able to directly challenge Canada's decision to revoke its patents through an appellate process.¹⁷⁵ In addition, Eli Lilly does not even contest that there was any procedural irregularity with the manner in which it was able to challenge the undesirable court decisions. However, Eli Lilly is now simply seeking another bite at the apple that would be unavailable to a domestic Canadian company. This is not within the traditional justification of investor-state disputes, which are supposed to provide a means for foreign investors that are otherwise without recourse. In fact, Eli Lilly's situation falls within one of the current criticisms of such disputes—that they unfairly provide more benefits to foreign investors than are available to domestic investors.

B. INVALIDATION OF PATENT RIGHTS SHOULD NOT CONSTITUTE EXPROPRIATION

Assuming that Eli Lilly has a covered investment, this Section explains why invalidation of Eli Lilly's patent rights should not constitute "expropriation" under the relevant investment chapter. This Section first demonstrates how Eli Lilly's case may involve a situation that is exempt from expropriation analysis. Alternatively, this Section examines how, although the situation may not be completely exempt based on prior decisions as well as policy grounds, the arbitration tribunal should not find that Canada has engaged in either direct or indirect expropriation.

Before addressing the specific legal claims, it is important to first clarify what expropriation means. All international agreements protecting foreign investments provide a claim against states that expropriate (take) investments covered by the agreement; such investments typically include not only tangible, but also intangible property of economic value. It is roughly analogous to U.S. takings law in terms of involving state action, but expropriations may exist in situations that would fail under U.S. takings law. In general, agreements recognize that there are some situations where countries should be allowed to expropriate investments, but only if there is a public purpose, the action is nondiscriminatory, and there is just compensation.

A foreign investor may have a claim for either direct or indirect expropriation. Direct expropriation claims involve outright and overt taking of property by the state, such as by transferring title to the state; the reason

175. In addition, although the Supreme Court of Canada denied Eli Lilly review, that was within the discretion of the Court.

for the taking is not important.¹⁷⁶ Indirect expropriation, on the other hand, can exist even if the investor maintains ownership of the investment, but because of “unreasonable interference” the investor loses all, or a significant part, of its investment.¹⁷⁷ Although the two types of expropriation are fairly straightforward to explain, there are few direct expropriation claims in recent times¹⁷⁸ and indirect expropriation analysis is complicated because tribunals use several different tests.

1. *Canada’s Actions May Be Exempt from an Expropriation Claim*

An initial question is whether there is any need to even address the details of an expropriation claim. There are two possible ways Canada could avoid the claim altogether. First, the situation could fall under a specific NAFTA intellectual property exception that prevents foreign investors from raising a claim for expropriation. Alternatively, the Canadian decisions may not constitute the “state action” that is a fundamental prerequisite to expropriation claims. Although there are arguments for excluding Eli Lilly’s claims under either of these grounds, a tribunal could reasonably find otherwise, as explained below.

First, Eli Lilly’s case is brought pursuant to NAFTA, which expressly excludes certain issues from consideration as expropriation that might otherwise qualify.¹⁷⁹ Article 1110 generally prohibits member states from expropriating foreign investments. However, paragraph seven states that it does not apply to “revocation, limitation or creation of intellectual property

176. For example, when Venezuela seized oil installations of foreign companies, it resulted in a series of direct expropriation cases. *E.g.*, P.G., *Venezuela and International Arbitration: Ick-SID*, *ECONOMIST* (Jan. 20, 2012, 6:03 PM), <http://www.economist.com/blogs/americasview/2012/01/venezuela-and-international-arbitration/>.

177. *E.g.*, August Reinisch, *Expropriation*, in *THE OXFORD HANDBOOK OF INTERNATIONAL INVESTMENT LAW* 407, 422 (Peter Muchlinski et al. eds., 2008). Accordingly, even if the investor continues to own legal title, there still may be indirect expropriation. *E.g.*, *Metalclad Corp. v. United Mexican States*, ICSID Case No. ARB(AF)/91/1, Award, ¶ 103 (Aug. 30, 2000), 5 ICSID Rep. 212 (2001) [hereinafter *Metalclad Award*].

178. *E.g.*, Anne K. Hoffmann, *Indirect Expropriation*, in *STANDARDS OF INVESTMENT PROTECTION* 151 (August Reinisch ed., 2008); Reinisch, *supra* note 177, at 408.

179. In addition, a similar analysis applies to many other existing and pending agreements that contain similar language about excluding as expropriation denial of intellectual property rights consistent with TRIPS. *E.g.*, U.S.-Singapore Free Trade Agreement, U.S.-Sing., art. 15.6.5, May 6, 2003; Trans-Pacific Partnership Agreement, Wikileaks Investment Chapter art. 11.7(5) (2015) [hereinafter draft TPP investment chapter]. Notably, TRIPS is a different agreement than NAFTA, but both have similar language concerning patentability requirements. Compare TRIPS, *supra* note 3, art. 27(1), with NAFTA, *supra* note 36, art. 1709(1).

rights” if consistent with the NAFTA provision on patents.¹⁸⁰ This seems to preclude expropriation claims of intellectual property, such as patents, that are revoked in a manner consistent with NAFTA.

In the Eli Lilly case, the issue is whether the revocation provision of NAFTA prevents countries from revoking patents because the interpretation of an existing patent law ground, utility, has changed since the patent was issued. Eli Lilly asserts that such a revocation is impermissible under NAFTA, which would mean that the paragraph seven exemption does not apply. However, NAFTA’s language does not explicitly support this conclusion and doing so would be contrary to recognized principles of how common law operates. There are two NAFTA sections on patents that are relevant to Eli Lilly’s situation. First, the most fundamental patent provision is article 1709(1), which requires each party to provide patents on inventions that satisfy the criteria of being new, useful, and non-obvious.¹⁸¹ Notably, NAFTA does not define what it means to be “useful,” so member states such as Canada should be able to define this as they wish, even if different than the laws of other NAFTA member states. Second, article 1709(8) states that countries may only revoke a patent when “grounds exist that would have justified a refusal to grant the patent.”¹⁸²

As noted earlier, courts do modify patent law standards and retroactively apply them. Given this reality, it seems reasonable to interpret the NAFTA “grounds exist” clause to mean that countries cannot revoke patents on a *new* ground that never previously existed, rather than modification of an existing ground. In Eli Lilly’s case, the patents were invalidated for failing to satisfy the utility doctrine—a long-existing ground for patentability. Although interpretation of the utility doctrine to incorporate the promise doctrine may be new, it is simply a modification of the existing ground. Nonetheless, it is unclear how a tribunal will in fact interpret this NAFTA provision, so it is not clear whether this exempts Canada’s actions from an expropriation claim.

The other issue is whether the Canadian decisions constitute state action, because state action is a fundamental requirement for expropriation. If Canada’s decisions are not a state action, then Eli Lilly has no expropriation claim. Unlike most investment arbitration cases where the action in question is a legislative or regulatory measure, Eli Lilly’s case involves solely the judiciary. Although there are only a handful of arbitration decisions involving domestic court actions, those decisions uniformly affirm that such actions

180. NAFTA, *supra* note 36, art. 1110(7).

181. *Id.* art. 1709(1).

182. *Id.* art. 1709(8).

can constitute state action.¹⁸³ Notably, even though actions of state courts *may* constitute state action, tribunals have stated this is only the case when the court ruling is clearly incompatible with a rule of international law, when there is a denial of justice, or when the state is responsible for a judicial decision “contrary to municipal law.”¹⁸⁴ There are no prior challenges to the substance of judicial decisions as expropriation. Rather, situations involved racial discrimination against an investor that a court failed to limit¹⁸⁵ as well as judicial interference with a contractually permitted arbitration.¹⁸⁶

The only possible basis for considering Canadian court actions against Eli Lilly to be state action is that those actions violate international law; Eli Lilly does not allege that the Canadian court decision was a denial of justice or that that judicial decision was contrary to municipal law. Eli Lilly claims that the judicial decision violated two separate international agreements—NAFTA and the PCT.

Eli Lilly alleges that the promise doctrine is inconsistent with NAFTA requirements concerning utility and nondiscrimination.¹⁸⁷ In particular, Eli Lilly asserts that a “dramatic and unanticipated shift” in Canada’s definition of utility is “significantly out of step” with its NAFTA partners.¹⁸⁸ However, that is irrelevant because NAFTA does not require member countries to have identical laws. Although NAFTA does require countries to grant patents that

183. *Azinian v. United Mexican States*, ICSID Case No. ARB(AF)/97/2, Award (Nov. 1, 1999), 14 ICSID REV.-FOR. INV. L.J. 538, 567 [hereinafter *Azinian Award*]; *Saipem S.p.A. v. People’s Republic of Bangl.*, ICSID Case No. ARB/05/7, Award, ¶¶ 189–90 (June 30, 2009) [hereinafter *Saipem Award*]; *see also Loewen Grp., Inc. v. United States*, ICSID Case No. ARB(AF)/98/3, Decision on Hearing of Respondent’s Objection to Competence and Jurisdiction, ¶ 70 (Jan. 5, 2001) [hereinafter *Loewen Decision*] (“The modern view is that conduct of an organ of the State shall be considered as an act of the State under international law, whether the organ be legislative, executive or judicial.”). However, these assertions are generally made in cases where *no* expropriation is found, and possibly cases where state action wasn’t even limited to the judiciary, such that they are dicta. *E.g.*, *Saipem Award*, *supra*, ¶ 191 (no expropriation found); *Azinian Award*, *supra*, ¶ 10 (state action was simply affirmation of city council decision); *Loewen Decision*, *supra*, ¶¶ 148, 241 (dismissing all claims both because claimant was not a qualifying investor and also because the claim was an attempt to use arbitration in lieu of a domestic appeal).

184. *Azinian Award*, *supra* note 183, ¶ 98; *Loewen Decision*, *supra* note 183, ¶ 47.

185. In *Loewen*, the investor claimed that racial and other inappropriate suggestions were made against it that resulted in the largest ever state verdict of over \$500 million for contracts worth less than \$5 million that when combined with a 125% bond requirement threatened to bankrupt the company, such that it could not realistically appeal. *E.g.*, Jake A. Baccari, *The Loewen Claim: A Creative Use of NAFTA’s Chapter 11*, 34 U. MIAMI INTER-AM. L. REV. 465, 468–69 (2003).

186. *Saipem Award*, *supra* note 183, ¶¶ 35–37, 39.

187. *Eli Lilly Notice of Arbitration*, *supra* note 2, ¶¶ 69–70.

188. *Id.* ¶ 9.

meet the standard of utility, it provides no definition, so nations are permitted to self-define it.¹⁸⁹ In addition, NAFTA does not state that countries are precluded from modifying its definition.

Eli Lilly also asserts that Canada has violated the NAFTA obligation to grant patents without discrimination as to field of technology.¹⁹⁰ In particular, Eli Lilly asserts that the promise doctrine has “almost exclusively” impacted pharmaceutical patents.¹⁹¹ However, the doctrine applies to all inventions, so this argument seems questionable. There are many facially neutral patent law standards that actually apply differently to different areas of technology.¹⁹² In addition, as explained by one WTO panel, a neutrally worded law does not de facto discriminate against a field of technology because it does not impact all fields equally.¹⁹³ In particular, the WTO panel stated that discrimination “does not prohibit bona fide exceptions to deal with problems that may only exist in certain product areas.” Such action is considered a permissible *differentiation*, rather than discrimination.¹⁹⁴

Eli Lilly also alleges that Canada’s action is inconsistent with NAFTA article 1709(8), which states that a country may revoke a patent only when “grounds exist that would have justified a refusal to grant the patent.” However, as noted earlier, this should be interpreted to mean that a country cannot revoke a patent on a new patentability requirement, but it should not bar a country from revoking a patent based on modification of the long-standing patent requirement of utility.

Eli Lilly claims that Canada’s modified utility requirement violates the international PCT rule barring countries from imposing “requirements as to the form or contents of the international application different from or additional to” those provided for in the PCT.¹⁹⁵ In particular, Eli Lilly asserts that the promise doctrine essentially requires certain information be disclosed in the patent application, and that these applications are a matter of form and content governed by the PCT for which a nation cannot make additional requirements.¹⁹⁶ However, as noted earlier, the PCT is an international

189. NAFTA, *supra* note 36, art. 1709(1).

190. *Id.*

191. Eli Lilly Notice of Arbitration, *supra* note 2, ¶ 69.

192. *E.g.*, Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575 (2003).

193. Panel Report, *Canada—Patent Protection of Pharmaceutical Products*, WT/DS114/R, ¶ 7.94 (Mar. 17, 2000).

194. *Id.* ¶ 7.92

195. Eli Lilly Notice of Arbitration, *supra* note 2, ¶ 45; Patent Cooperation Treaty, *supra* note 51, art. 27(1).

196. Eli Lilly Notice of Arbitration, *supra* note 2, ¶ 46.

agreement intended to simplify patent filings on a global basis without restricting *substantive* patentability conditions in individual countries, such as utility. However, even with respect to disclosures in the application, there is prior precedent for nations requiring additional disclosures beyond what is in the PCT. For example, the United States requires that patent applicants disclose best mode in the patent application, even though that is not a requirement of the PCT.¹⁹⁷

Accordingly, there are reasonable arguments for considering Canada's actions as not expropriation. Nonetheless, the explicit exemption from NAFTA expropriation claims involving "revocation" of intellectual property rights consistent with the separate NAFTA section on patents is sufficiently ambiguous, such that it is not clear how a tribunal would rule, even though this author believes Eli Lilly should be within this exclusion. Similarly, a tribunal could find for the first time that a domestic court decision modifying existing common law is a state action based on a violation of international law, even though this author believes that is incorrect.

2. *Canada Could Be Found to Not Have Expropriated Eli Lilly's Patents*

Although there are legitimate reasons why a tribunal should exclude Eli Lilly's case from an expropriation analysis as noted in the above section, this section will consider whether Eli Lilly has expropriation claims based on traditional expropriation concepts since it is unclear how a tribunal would rule. Eli Lilly has alleged that Canada directly and indirectly expropriated its patent rights in an unusual case that is not typical of either claim. As explained below, a tribunal should find that Canada committed neither type of expropriation under NAFTA's investment chapter. Although this Section analyzes both types, there is substantially more to analyze with indirect expropriation claims because there are several independent tests that tribunals use.

a) *Canada Should Not Be Found to Have Directly Expropriated Eli Lilly's Patents*

The first question is whether Canada directly expropriated Eli Lilly's investment. Canada did remove Eli Lilly's title to previously granted patents, which is typical of direct, rather than indirect expropriation claims. However, unlike in most direct expropriation claims, ownership of those patent rights were not transferred to Canada or any other party; rather, what was in those patents is now in the public domain and freely useable by anyone. There is a possible argument that patent invalidation is tantamount to physical property

197. 35 U.S.C. § 112(1) (2012).

seized by the state in terms of the benefit to the state—even though benefit is not necessary for direct expropriation. Similar to the situation where direct expropriation of tangible property would benefit the state, invalidation of Eli Lilly's property rights arguably benefits all Canadian citizens. For example, other companies can now make and sell generic versions of Eli Lilly's drugs in Canada because there is no valid patent to bar them; this obviously benefits these companies. Moreover, since such companies are likely generic companies that compete based on price, Canadian citizens benefit from the lower cost of drugs. Nonetheless, since benefit is not required for direct expropriation, this argument is weak. The important issue here is that unlike direct expropriation cases, the titles to Eli Lilly's investments—its patents—were not transferred to the state since no one owns them at all.

b) Canada Should Not Be Found to Have Indirectly Expropriated Eli Lilly's Patents

The next issue is whether Canada has committed indirect expropriation through invalidation of Eli Lilly's patents. Usually, indirect expropriation claims mean that the investor retains title, but there is “unreasonable interference” as well as “deprivation” of property rights, such that the investor loses all, or a significant part of its investment.¹⁹⁸ This theory seems to better fit Eli Lilly's case. Eli Lilly could assert that while it technically still owns the patents at issue, they have no economic value because without valid patents, Eli Lilly cannot charge a premium price because there will be other competitors. However, as will be shown below, the tribunal should not find that Canada indirectly expropriated Eli Lilly's investments.

An important issue is how to analyze indirect expropriation. Although many agreements, including NAFTA do not provide criteria for evaluating indirect expropriation, there are two basic approaches.¹⁹⁹ First, indirect expropriation may exist based solely on the effect of the interference with the investment, such that it is called the “sole effect doctrine.” However, many tribunals and scholars consider this approach unfair and instead weigh economic impact on an investment against other factors including legitimate state interest, proportionality between state interest and investor harm, as well as reasonable expectations.²⁰⁰ As explained below, Eli Lilly's stronger

198. Reinisch, *supra* note 177, at 422. Accordingly, even if the investor continues to own legal title, there still may be indirect expropriation. *E.g.*, Metalclad Award, *supra* note 177, ¶ 103.

199. *E.g.*, NAFTA37, art. 1110; *see also* SUSY H. NIKIÈMA, INT'L INST. FOR SUSTAINABLE DEV., BEST PRACTICES: INDIRECT EXPROPRIATION 5 (2012).

200. Andrew Newcombe, *The Boundaries of Regulatory Expropriation in International Law*, 20 ICSID REV.-FOR. INV. L.J. 1, 9–11 (2005); NIKIÈMA, *supra* note 199, at 13. In addition, some

claim is under the sole effect doctrine, but utilizing that doctrine seems fundamentally unfair.

i) Sole Effect Doctrine Favors Lilly But Should Not Be Applied

Under the sole effect doctrine, significant and irreversible damage to enjoyment of property is the sole criterion for finding indirect expropriation.²⁰¹ Generally, tribunals speak of damage that is so severe that there is no longer any economic interest to the investor; for example, one tribunal stated that rights must be “rendered so useless that they must be deemed to have been expropriated.”²⁰² Accordingly, economic activity that is made more difficult, but not impossible will likely not constitute indirect expropriation. The state intent or possible benefit is not relevant pursuant to this doctrine.²⁰³

If the sole effect test is applied, Eli Lilly seems to have a strong claim. Invalidation of a patent is an absolute and permanent interference since the patent owner has no rights after it is invalidated.²⁰⁴ Indeed, prior commentators have noted that actions short of invalidation of patents would meet this standard, such as a compulsory license of a patent in which the

have noted that the sole effect doctrine has been primarily recognized until recently when wealthy countries have become more subject to investor-state disputes and thus interested in emphasizing legitimate state interest. *E.g.*, SEBASTIÁN LÓPEZ ESCARCENA, *INDIRECT EXPROPRIATION IN INTERNATIONAL LAW* 10 (2014).

201. *E.g.*, *Consortium R.F.C.C. v. Kingdom of Morocco*, ICSID Case No. ARB/00/6, Award, ¶ 68 (Dec. 22, 2003), 20 ICSID REV.-FOR. INV. L.J. 391 (2005).

202. *Starrett Hous. Corp. v. Iran*, Interlocutory Award No. ITL 32-24-1, 4 Iran-US Cl. Trib. Rep. 122, 154 (Dec. 19, 1983); *see also* *Pope & Talbot Inc. v. Canada*, Interim Award, ¶ 102 (NAFTA Arb. Trib. June 26, 2000), 40 I.L.M. 258 (2001) (considering whether state interference is “sufficiently restrictive to support a conclusion that the property has been taken from the owner”); *Metalclad Award*, *supra* note 177, ¶ 103 (requiring action that “has the effect of depriving the owner, in whole or in significant part, of the use or reasonably-to-be-expected economic benefit of property”); *Técnicas Medioambientales Tecmed S. A. v. United Mexican States*, ICSID Case No. ARB(AF)/00/2, Award, ¶ 115 (May 29, 2003), 10 ICSID Rep. 134 (2006) [hereinafter *Tecmed Award*] (“[R]adically deprived of the economical use and enjoyment of its investments, as if the rights related thereto . . . had ceased to exist.”).

203. *E.g.*, *Newcombe*, *supra* note 200, at 11–12; *Metalclad Award*, *supra* note 177, ¶ 103 (asserting no need to consider “motivation or intent” of state action because indirect expropriation can exist “even if not necessarily to the benefit of the host state”); *see also* *Tippetts v. TAMS-AFFA Consulting Eng'rs of Iran*, Award No. 141-7-2, 6 Iran-US Cl. Trib. Rep. 219, 225–26 (June 29, 1984) (government intention is less important than effect of measure on owner of assets).

204. Once a patent is invalidated, there is no prospect for obtaining a new patent because the original patent would preclude a subsequent application for the same invention from satisfying the requirement of being “new.” *See* 35 U.S.C. § 102(b) (2012).

patent exists, but the ability to exclusively determine how to exploit it is limited.²⁰⁵

However, a number of commentators and tribunals in recent years have suggested that the sole effect doctrine is unfair and inappropriate. Although the sole effect test was the primary test applied by tribunals since the 1980s and even through the early 2000s, more recent tribunals have shifted away from this doctrine.²⁰⁶ In particular, tribunals have shifted toward weighing harm to the investment against the state interest. In addition, although typically an element in fair and equitable treatment standards, tribunals are also increasingly incorporating legitimate expectations of investors into their analysis of indirect expropriation claims.²⁰⁷ Recent agreements as well as model agreements tend to explicitly enumerate these as factors for consideration, which notably mirror the factors that the U.S. Supreme Court utilizes to determine whether there has been a regulatory taking.²⁰⁸ Accordingly, analyzing Eli Lilly's claims pursuant to these factors seems to represent strong policy. As this section will explain, Eli Lilly has a much weaker claim when these factors are considered.

205. E.g., Christopher Gibson, *A Look at the Compulsory License in Investment Arbitration: The Case of Indirect Expropriation*, 25 AM. U. INT'L L. REV. 357, 386 (2010); Tsai-Yu Lin, *Compulsory Licenses for Access to Medicines, Expropriation and Investor-State Arbitration Under Bilateral Investment Agreements—Are there Issues Beyond the TRIPS Agreement?*, 40 IIC: INT'L REV. INTELL. PROP & COMPETITION L. 152 (2009).

206. E.g., Ursula Kriebaum, *Expropriation*, in INTERNATIONAL INVESTMENT LAW 38–41 (Marc Bungenberg et al. eds., 2015), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2295979.

207. E.g., DOLZER & SCHREUER, *supra* note 65, at 115–17; see also 2004 Canadian Model BIT Annex B 13(1)(b)(ii) (expressly noting the “extent to which the measure or series of measures interfere with distinct, reasonable investment-backed expectations” is a factor that should be considered in whether there is an indirect expropriation).

208. E.g., US Model BIT Annex B, Expropriation; Agreement on the Encouragement and Reciprocal Protection of Investments, U.S.-Uruguay., Annex B, para 4(a), Nov. 5, 2005, T.I.A.S. No. 06-1101 (considering economic impact of government action, the extent to which the government action interferes with distinct, reasonable, investment based expectations, and the character of the government action); 2004 Canadian Model BIT Annex B 13(1)(b); ASEAN Comprehensive Investment Agreement of 2009, Annex 2 (3) (considering whether a binding written commitment by the government has been breached and also considering the character of the government action and whether it is disproportionate to its public purpose). Importantly this specifically rejects the sole effect doctrine by stating that adverse effect on the economic value of an investment is not on its own adequate to establish indirect expropriation. *Id.*

ii) Legitimate State Interest Should Outweigh Eli Lilly's Interest

Although all expropriations must be for a public purpose, considering the purpose behind the state action is nonetheless important for two possible reasons. Some tribunals consider that when a state action is pursuant to its regulatory police powers, there should be no compensable expropriation, so long as the action is done on a nondiscriminatory basis and pursuant to due process.²⁰⁹ Even for tribunals that do not completely exclude state action from the scope of compensable expropriation, the type of state interest is relevant in considering whether the state action is proportional to investor harm.²¹⁰

An important issue is what constitutes a legitimate interest of the state. Although this is often considered to be synonymous with regulatory police powers, there is no internationally agreed definition of such powers.²¹¹ Nonetheless, the 2012 US Model BIT explicitly noted that legitimate public welfare objectives that would usually not constitute indirect expropriation include public health, safety, and the environment.²¹² This is also consistent with prior tribunal decisions, such as *Methanex*, in which the tribunal found that a law barring use of a petrol additive deemed carcinogenic was a bona fide regulation that served legitimate public interest, such that it was not compensable.²¹³

Countries have strong policy interests in limiting the scope of intellectual property rights to situations where the rights result in more benefits than harm. As noted earlier, it would be unfair to impose the economic cost of higher prices attendant with patent protection unless the inventor of the patent provided an adequate exchange through proper disclosure of the invention. The promise doctrine that Eli Lilly challenges aims to ensure that this fundamental patent bargain is satisfied.

209. *E.g.*, *Saluka Invs. B.V. v. Czech Republic*, Partial Award, ¶ 262 (Perm. Ct. Arb. Mar. 17, 2006) [hereinafter *Saluka Partial Award*]; *Methanex Corp. v. United States*, Final Award of the Tribunal on Jurisdiction and Merits, Part IV.D, ¶ 7 (NAFTA/UNCITRAL Arb. Trib. Aug. 3, 2005), 16 ICSID Rep. 40 (2012) [hereinafter *Methanex Final Award*]; *Cont'l Cas. Co. v. Arg. Republic*, ICSID Case No. ARB/03/9, Award, ¶ 276 (Sept. 5, 2008) [hereinafter *Cont'l Cas. Co. Award*]; *Feldman v. United Mexican States*, ICSID Case No. ARB(AF)/99/1, Award, ¶ 105 (Dec. 16, 2002), 7 ICSID Rep. 341 (2005); *S.D. Myers Partial Award*, *supra* note 135, ¶ 281.

210. *E.g.*, *Tecmed Award*, *supra* note 202, ¶ 122.

211. *E.g.*, *Saluka Partial Award*, *supra* note 209, ¶ 263 (“[I]nternational law has yet to identify in a comprehensive and definitive fashion precisely what regulations are considered . . . within the police or regulatory power of States.”).

212. US Model BIT, Annex B, Expropriation, ¶ 4(b).

213. *Methanex Final Award*, *supra* note 209, at Section IV.D, ¶ 15.

Although this author thinks there is a clear case for considering the design of intellectual property rights to be a legitimate state interest, this is admittedly different than traditionally listed public welfare objectives. The closest common public welfare objective is the state interest in promoting public health. Arguably, this is promoted by denying patents on drugs that would increase the cost of medicine and thereby negatively impact public health for those could not afford the drugs. However, it is unclear if a tribunal would agree. Indeed, many would suggest that denying patents on drugs to reduce the cost of medicine is poor policy since that would reduce the incentive to create new drugs. Basically, the argument is that expensive drugs for the short-term are better than no new drugs in the long term. On the other hand, the proper balance between incentivizing new drugs that are expensive to consumers and potentially also impede subsequent research is a thorny question to which there are no uniform answers. This is fundamentally a policy determination that nations previously were able to decide based on their domestic preferences for promoting innovation versus access prior to TRIPS. Importantly, it is not simply an issue of developed versus developing countries; some developed countries, such as Italy and Portugal only granted patents on drug patents recently. In addition, before Canada concluded NAFTA, which has similar patent requirements as TRIPS, it granted drug patents, but permitted them to be broadly licensed by generic companies. Although the United State has traditionally provided expansive scope of patentability with little regard for impacts on access, other countries bar patents that reduce access to treatment, such as patents on methods of medical treatment.²¹⁴ Given that there is broad disagreement concerning whether there is a public health interest in reducing the cost of medicine by denying some patents on drugs, it is unclear how a tribunal would rule.

Assuming that Canada has a legitimate interest in tailoring its patent laws to best promote access to affordable medicine while consistent with international law, the next step is to consider whether that interest unduly harms Eli Lilly's investment. Some tribunals are deferential to self-declared state interests and find no expropriation so long as the state action is nondiscriminatory and in accordance with due process.²¹⁵ However, other tribunals apply a proportionality test, balancing the public purpose against the investor's expectations. This can be tricky because although a balancing test is more reasonable than the sole effect doctrine, it depends on how a

214. *E.g.*, European Patent Convention, *supra* note 42, art. 53(c); *see also* TRIPS, *supra* note 3, art 27(3) (permitting exclusion of such inventions from the scope of patentability).

215. *E.g.*, Chemtura Corp. v. Canada, Award, ¶ 266 (NAFTA/UNCITRAL Arb. Trib. Aug. 2, 2010).

tribunal applies this standard. For example, in *Tecmed*, the tribunal found that even if there is a valid state interest, it cannot outweigh the investor interest unless the state action is necessary to achieve the intended public interest, which it defined as the only measure available to achieve the objective, or the least detrimental among a number of effective solutions.²¹⁶ In that case, the tribunal found that the state's refusal to renew a license for a hazardous waste treatment plant was indirect expropriation because even though the license was denied for the legitimate interest of resolving local complaints concerning health and safety, there were less detrimental solutions possible, such as relocation of the plant.²¹⁷

Eli Lilly's situation seems somewhat similar to *Tecmed's* in that its entire interest (in its patents) was vitiated when there arguably could have been a less detrimental solution. Just as the *Tecmed* tribunal suggested that the state could have taken a different action that would not have entirely terminated the investor's interest, so too Canada's law may seem unduly severe. In particular, although the policy reason for Canadian law is well established, given that Canada is the only country to have this law, a tribunal could find that it may not be necessary to apply the law in this manner.²¹⁸ However, some commentators note that although Canada is the only country to consider patent promises with respect to the utility requirement, other countries, including Australia and New Zealand have similarly invalidated patents that fail to achieve what they promise, despite having some basic utility.²¹⁹

iii) Eli Lilly Has No Legitimate Investor Expectations That Have Been Violated

The best consideration for rejecting Eli Lilly's claim for indirect expropriation is that it did not have any legitimate expectations that were violated. Recent agreements specifically note analysis of legitimate expectations of an investor as a way to assess indirect expropriation, but tribunals have also considered legitimate expectations pursuant to agreements that do not explicitly require this method of assessment. In both cases, the

216. *Tecmed Award*, *supra* note 202, ¶ 122.

217. *Id.* ¶ 51.

218. This is particularly true if the tribunal is sympathetic to Eli Lilly's position, even though there are competing expert reports about whether Canada's law is in fact unique. *Compare* Memorial of Claimant Eli Lilly, ¶¶ 145–60 (citing experts that allege that Canada's law is unique) *with* Counter Memorial of Canada, ¶¶ 170–99 (noting lack of international harmony in standards and functional similarity between US and Canadian law).

219. *E.g.*, *Pracdes Pty. Ltd. v. Stanilite Electronics Pty. Ltd.* (1995) 35 IPR 259, 273–75 (Austl.); *Hammar Maskin A.B. v Steelbro N.Z. Ltd.* [2010] NZCA 83, ¶ 76 (Austl.).

focus is on whether the investor had a legitimate expectation that was violated. The legitimacy of the investor's expectations generally depends on specific assurances by the state.²²⁰ This is consistent with the interpretation of legitimate expectations in the context of violation of the standard of "fair and equitable treatment" from which the concept is derived.²²¹ The facts of prior cases may help to put this standard in context.

Metalclad v. United Mexican States provides a helpful example of legitimate investor expectations to support an indirect expropriation claim. In that case, there were multiple assurances that were specifically relied upon in making investments.²²² Investor Metalclad obtained a state construction permit for a hazardous waste landfill and assurances from federal agents that all necessary permits had been required.²²³ However, after construction begun, the local government ordered construction to stop because Metalclad had not obtained a municipal construction permit.²²⁴ At that point, federal agents once again assured Metalclad that if it applied for such a permit, it would be granted.²²⁵ However, it was not²²⁶ and thus Metalclad could not operate the landfill and the state later declared the land a Natural Area for protection of a rare cactus thereby terminating any possibility of Metalclad operating its facilities.²²⁷ The tribunal found that Metalclad had reasonably relied on assurances by the federal government and had a reasonable expectation to construct and operate the landfill that was thwarted.²²⁸

In contrast, in a more recent decision, a different NAFTA tribunal took a different approach to indirect expropriations in *Methanex v. United States*. In that case, the Canadian methanol producer asserted that a California law banning a carcinogenic gasoline additive resulted in an indirect expropriation of its investments in the California and U.S. market and improperly benefited

220. *E.g.*, Reinisch, *supra* note 177, at 448. This is also consistent with the fair and equitable treatment standard from which the concept of legitimate expectations are derived.

221. *E.g.*, Katia Yanaca-Small, *Fair and Equitable Treatment Standard: Recent Developments*, in STANDARDS OF INVESTMENT PROTECTION 111, 126 (Reinisch, ed. 2008) ("legitimate expectation is assumed more readily if an individual investor receives specific formal assurances" from the government official that the official should perceive the investor to reasonably rely upon); CAMPBELL MCLACHLAN, LAURENCE SHORE & MATTHEW WEINIGER, INTERNATIONAL INVESTMENT ARBITRATION: SUBSTANTIVE PRINCIPLES ¶ 7.99 (2007).

222. Metalclad Award, *supra* note 177, ¶¶ 37–44.

223. *Id.*

224. *Id.* ¶ 40.

225. *Id.* ¶ 41.

226. *Id.* ¶¶ 45–50.

227. *Id.* ¶¶ 45, 69.

228. *Id.* ¶ 104.

the domestic ethanol industry.²²⁹ Although the tribunal dismissed the case on other grounds,²³⁰ it made findings that a state measure prohibiting use of a petrol additive considered carcinogenic was not an indirect expropriation because it was a bona fide regulation to serve the public interest on a nondiscriminatory basis and that there were no specific state representations to induce the investor.²³¹ In particular, the tribunal found that the investor should not have been surprised that environmental and health protection laws might change and adversely impact its interests.²³²

If Eli Lilly's legitimate expectations were considered, there should be no expropriation because Eli Lilly was given no specific assurance that either the law would not change, or that its patent would remain forever valid. Although there is a presumption of validity for issued Canadian patents, the fact that it is a presumption, rather than an ironclad right, suggests that there is no reasonable expectation that it will be immune from cancelation.²³³ Indeed, patent scholars have previously noted that the public should not expect issued patents to be valid based on the current system which in fact relies on litigation challenges as a more efficient mechanism to weed out improper patents, than to have patent offices spend more time preventing invalid patents from issuing.²³⁴ Although Eli Lilly complains that it was

229. Methanex Final Award, *supra* note 209, Part III.A, ¶ 1, Section IV.D, ¶ 2.

230. *Id.* at Section IV.F, ¶ 1.

231. *Id.* at Section IV.D, ¶ 7.

232. *Id.* at Section IV.D, ¶ 10; *see also* Parkerings-Compagniet A.S. v. Republic of Lith., ICSID Case No. ARB/05/8, Final Award, ¶ 331–35 (Sept. 11, 2007) [hereinafter Parkerings Final Award] (affirming that explicit promise or assurance is necessary and without that investor's assertion that changes to domestic law were unfair did not constitute violation of fair and equitable treatment, especially given that investor knew that the country was in a state of transition from being part of the Soviet Union to part of the EU, such that changes were likely).

233. This is underscored not only by the fact that Canadian patent law permits issued patents to be challenged (subject to the presumption of validity), but also provides mechanisms for such patents to be challenged.

234. *E.g.*, Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. (2001). *But see* F. Scott Kieff, *The Case for Preferring Patent-Validity Litigation Over Second-Window Review and Gold-Plated Patents: When One Size Doesn't Fit All, How Could Two Do the Trick?*, 157 U. PENN. L. REV. 1937 (2009). In addition, although there have been proposals to more closely scrutinize certain patent applications where the inventor desires a stronger presumption of validity, no such proposals have been enacted. *E.g.*, Alan Devlin, *Revisiting the Presumption of Patent Validity*, 37 SW. U. L. REV. 323 (2008); Mark A. Lemley et al., *What to Do About Bad Patents*, REG., Winter 2005–06, at 10; Vincent M. De Grandpré, *United States: U.S. Looks at 'Gold-plating' to Enhance Patents*, MONDAQ, <http://www.mondaq.com/unitedstates/x/70826/Patent/US+Looks+At+GoldPlating+To+Enhance+Patents/> (last updated Dec. 2, 2008) (noting that President Obama proposed this during his presidential campaign). *But see* F. Scott Kieff, *supra* (arguing against such proposals).

shocked by Canada's change in the law, this is an inadequate ground for a claim of legitimate investor expectations given that Canada provided no specific assurance to Eli Lilly. In addition, although Eli Lilly seems to believe that an issued patent should be considered an assurance that the patent will remain valid, Eli Lilly's assumption is fundamentally inconsistent with patent law in Canada and other countries. As noted earlier, an issued patent is only presumptively valid, but can be and often is subsequently invalidated if it is later found not to meet patentability requirements. In addition, the mere grant of a patent seems very different than the multiple assurances given to the investor that the investor then relied upon to its detriment in *Metalclad*.²³⁵ Whereas the investor in *Metalclad* expended funds in building a hazardous landfill in reliance on the multiple investments, there is no claim that Eli Lilly developed its drugs in reliance on Canadian law. To the contrary, Eli Lilly developed its drugs as any multinational pharmaceutical company does—to sell worldwide.

C. INVALIDATION OF PATENT RIGHTS SHOULD NOT CONSTITUTE A VIOLATION OF THE FAIR AND EQUITABLE TREATMENT STANDARD

Assuming that Eli Lilly has an appropriate “investment” under NAFTA, even if it cannot establish an expropriation claim, it can alternatively recover compensation if Canada failed to provide “fair and equitable treatment” to Eli Lilly's investment.²³⁶ This claim is very important to Eli Lilly and all other investors since tribunals tend to take a flexible interpretation of what constitutes “fair and equitable treatment.” Claims based on a violation of the fair and equitable treatment standard are currently the most common²³⁷ and successful type of investment claim,²³⁸ and often prevail even when there is

235. *Metalclad Award*, *supra* note 177, ¶ 4108.

236. NAFTA, *supra* note 36, art. 1105(1).

237. UNCTAD, FAIR AND EQUITABLE TREATMENT: UNCTAD SERIES ON ISSUES IN INTERNATIONAL INVESTMENT AGREEMENTS II, A SEQUEL 10 (2012) [hereinafter UNCTAD, FAIR AND EQUITABLE TREATMENT]; Rudolf Dolzer, *Fair and Equitable Treatment: Today's Contours*, 12 SANTA CLARA J. INT'L L. 7, 10 (2013); *see also* Mercurio, *supra* note 26, at 894 (noting that although the standard was traditionally “rarely invoked” and only applicable where action was “egregious and shocking,” it is now commonly invoked due to a significantly broadened interpretation of the standard since the early 2000s). Most BITs and trade agreements include such standards, although a few BITs with Asian countries do not. Katia Yannaca-Small, *Fair and Equitable Treatment Standard: Recent Developments*, in STANDARDS OF INVESTMENT PROTECTION 110, 113 (August Reinisch ed., 2008).

238. In 2012 alone, six of the twelve published decisions finding state liability did so based on a violation of fair and equitable treatment. Violation of the fair and equitable treatment standard was the most common ground for state liability. UNCTAD, RECENT DEVELOPMENTS, *supra* note 15, at 5. A prescient professor noted in 1981 that “the right to fair and equitable treatment goes much further than the right to most favored-nation and to

no indirect expropriation.²³⁹ Nonetheless, this section explains why a tribunal should find that Eli Lilly does not have a valid claim against Canada for violation of the fair and equitable treatment standard because Eli Lilly had no legitimate expectations that were violated, which is the crux of this standard.

A key question is what constitutes “fair and equitable treatment.” Technically, there are differences in treaty language governing foreign investments that use the phrase “fair and equitable treatment.” Some, such as NAFTA, link the phrase to only minimum standards of conduct pursuant to customary international law,²⁴⁰ whereas others have no reference for what constitutes “fair and equitable treatment.”²⁴¹ However, in practice, tribunals seem to treat all claims similarly.²⁴² Essentially, tribunals as well as scholars consider whether there is a violation based on a number of factors.²⁴³ These include (a) defeating investors’ legitimate expectations (sometimes in balance

national treatment . . . so general a provision is likely to be almost sufficient to cover all conceivable cases and it may well be that provisions of the Agreements affording substantive protection are not more than examples of specific instances of this overriding duty.” F. A. Mann, Note, *British Treaties for the Promotion and Protection of Investments*, 52 BRIT. Y.B. INT’L L. 241 (1981).

239. E.g., Matthew C. Porterfield, *State Practice and the (Purported) Obligation Under Customary International Law to Provide Compensation for Regulatory Expropriations*, 37 N.C. J. INT’L L. & COM. REG. 159, 168–70 (2011); Marcela Klein Bronfman, *Fair and Equitable Treatment: An Evolving Standard*, 10 MAX PLANCK Y.B. UNITED NATIONS L. 609, 648 (2006); Yannaca-Small, *supra* note 237, at 112.

240. NAFTA, *supra* note 36, art. 1105(1). This should technically limit violations to situations where the action is “shocking, egregious and outrageous.” E.g., Jack J. Coe, *Fair and Equitable Treatment Under NAFTA’s Investment Chapter*, 96 ASIL PROC. 17 (2002).

241. E.g., Free Trade Agreement, China-Switz., art. 4, July 6, 2013, *available at* <http://www.seco.admin.ch/themen/00513/00515/01330/05115/index.html>; Agreement on the Reciprocal Promotion and Protection of Investments, BLEU-Taj., art. 3, Feb. 12, 2009, *available at* <http://investmentpolicyhub.unctad.org/Download/TreatyFile/418>; *see also* MAHNAZ MALIK, IISD, BEST PRACTICES SERIES BULLETIN: FAIR AND EQUITABLE TREATMENT 3 (Sept. 2009); UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at 6–7.

242. Notably, even where tribunals are interpreting the standard pursuant to an agreement that requires the standard to be linked to customary international law, panels do not necessarily do so and may instead simply rely on other tribunal decisions that do not require consideration of international law. E.g., R.R. Dev. Corp. v. Republic of Guad., ICSID Case No. ARB/07/23, Award, ¶ 219 (June 29, 2012); *see also* UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at xv, 11; Matthew C. Porterfield, *A Distinction Without a Difference? The Interpretation of Fair and Equitable Treatment Under Customary International Law by Investment Tribunals*, INVESTMENT TREATY NEWS (Mar. 22, 2013), <http://www.iisd.org/itn/2013/03/22/a-distinction-without-a-difference-the-interpretation-of-fair-and-equitable-treatment-under-customary-international-law-by-investment-tribunals/>.

243. Alternatively, some suggest that there is no stable or fixed content to this standard. E.g., IOANA TUDOR, THE FAIR AND EQUITABLE TREATMENT STANDARD IN THE INTERNATIONAL LAW OF FOREIGN INVESTMENT 133 (2008).

with the host state's right to regulate), (b) denial of justice and due process, (c) manifest arbitrariness in decision making, (d) undue discrimination, or (e) outright abusive treatment.²⁴⁴ Not all of these factors need be present in every case, but legitimate investor expectations are considered key to establishing a violation.²⁴⁵

For that reason, the only relevant factor to consider is whether Eli Lilly had a legitimate expectation that was defeated.²⁴⁶ In particular, the issue is whether Canada unexpectedly changed its law such that Eli Lilly's legitimate expectations when it made its investment were violated. Obviously, the critical question is what constitutes "legitimate expectations."

244. *E.g.*, UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at 62; TUDOR, *supra* note 243, at 155; Barnali Choudhury, *Evolution or Devolution?: Defining Fair and Equitable Treatment in International Investment Law*, 6 J. WORLD INVESTMENT & TRADE 297, 302–15 (2005); Yannaca-Small, *supra* note 237, at 129. Alternatively, panels cite a quotation from *Waste Management v. Mexico* that addresses similar factors. *E.g.*, GAMI Invs., Inc. v. United Mexican States, Final Award, ¶ 89 (NAFTA/UNCITRAL Arb. Trib. Nov. 15, 2004) [hereinafter GAMI Final Award]; Methanex Final Award, *supra* note 209, at Part IV.C, ¶ 26; Siemens A.G. v. Arg. Republic, ICSID Case No. ARB/02/8, Award, ¶ 297 (Feb. 6, 2007); Azurix Corp. v. Arg. Republic, ICSID Case No. ARB/01/12, Award, ¶ 370 (July 14, 2006); Waste Mgmt., Inc. v. United Mexican States, ICSID Case No. ARB(AF)/00/3, Award, ¶ 98 (Apr. 30, 2004).

245. *E.g.*, *Electrabel S.A. v. Republic of Hung.*, ICSID Case No. ARB/07/19, Decision on Jurisdiction, ¶ 7.75 (Nov. 30, 2012) ("most important function" of standard is to protect legitimate expectations); *Saluka Partial Award*, *supra* note 209, ¶ 302 ("[T]he standard of fair and equitable treatment is . . . closely tied to the notion of legitimate expectations which is the dominant element of the standard."); *EDF (Servs.) Ltd. v. Romania*, ICSID Case No. ARB/05/13, Award, ¶ 216 ("one of the major components of the [fair and equitable treatment] standard is the parties' legitimate and reasonable expectations") [hereinafter *EDF (Servs.) Ltd.*]; see also Dolzer, *supra* note 237, at 17 (noting that "protection of legitimate expectations . . . 'is the central pillar' of the standard).

246. At first glance, "manifest arbitrariness" or "discrimination" may seem relevant to Eli Lilly's claim that Canada breached its obligation to "refrain from conduct that is arbitrary, unfair, unjust and discriminatory" in invalidating its two patents. Eli Lilly Notice of Arbitration, *supra* note 2, ¶ 81. However, this is unlikely since manifest arbitrariness without direct targeting of a foreign investor requires an act that shocks judicial propriety and cannot even include a country failing to follow its own laws. *Case Concerning Elettronica Sicula S.p.A. (ELSI) (U.S. v. It.)*, Judgment, 1989 I.C.J. 15, ¶ 128, (July 20) (requiring conduct that "shocks, or at least surprises, a sense of juridical propriety"); *Cargill, Inc. v. United Mexican States*, ICSID Case No. ARB(AF)/05/2, Award, ¶ 303 (Sept. 18, 2009) (finding manifestly arbitrary conduct where Mexico imposed an import permit for high fructose corn syrup with the express intent of damaging U.S. producers of such syrup and where there were no objective criteria for how to obtain such permits). Similarly, *undue* discrimination generally requires treating an investor differently because of impermissible categories such as race and gender, or at a minimum, treating the investor differently than domestic investors. UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at 82. Eli Lilly has not made any such argument.

Although agreements generally do not define “legitimate expectations,” there are essentially two approaches.²⁴⁷ The broadest and most investor-friendly approach is that a state must ensure a stable legal and business environment.²⁴⁸ The other approach only finds legitimate expectations if those expectations arise from a specific state representation that the investor relies on.²⁴⁹ Indeed, one tribunal cautioned against an unduly broad reading of legitimate expectations that would inappropriately constrain states.²⁵⁰ Accordingly, a number of tribunals and scholars suggest not only that there should be a specific state representation, but also that state interests should be balanced against investor expectations.²⁵¹ This author believes that the standard grounded in state representation is preferable as a matter of policy. Other commentators also endorse this balancing approach, and arguably tribunals have recently trended toward it. In particular, commentators as well as some tribunals have noted that it is difficult to find and justify legitimate expectations solely on the basis of a preexisting legal regime and that this standard should not serve the same purpose as stabilization clauses.²⁵² Nonetheless, both standards are discussed below to predict how a tribunal

247. Arguably, there are other variations of legitimate expectations, but the two approaches outlined here represent the extremes.

248. *E.g.*, UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at 64; Dolzer, *supra* note 237, at 24–25.

249. *E.g.*, UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at 67–70; Dolzer, *supra* note 237, at 24–25.

250. EDF (Servs.) Ltd., *supra* note 245, ¶ 217 (“The idea that legitimate expectations, and therefore [fair and equitable treatment], imply the stability of the legal and business framework, may not be correct if stated in an overly broad and unqualified formulation. The [fair and equitable treatment] might then mean the virtual freezing of the legal regulation of economic activities, in contrast with the State’s normal regulatory power and the evolutionary character of economic life.”).

251. After all, although agreements aim to protect investments, investors arguably benefit the host state, such that public policy concerns of the host state should be considered. *E.g.*, Abhijit P.G. Pandaya & Andy Moody, *Legitimate Expectations in Investment Treaty Arbitration: An Unclear Future*, 15 TILBURG L. REV. 93, 96 (2010–11). In addition, some suggest that in applying this standard, tribunals should be deferential to host states.

252. *E.g.*, CAMPBELL MCLACHLAN ET AL., INTERNATIONAL INVESTMENT ARBITRATION: SUBSTANTIVE PRINCIPLES 238 (2008); Pandaya & Moody, *supra* note 251, at 114; EDF (Servs.) Ltd., *supra* note 245, ¶ 218 (fair and equitable treatment should not serve same function as stabilization clauses); Michele Potestà, *Legitimate Expectations in Investment Treaty Law: Understanding the Roots and the Limits of a Controversial Concept*, 28 ICSID REV. 88, 114 (2013) (noting that it would be “illogical” to permit fair and equitable treatment to provide the same type of protection as a stabilization clause that an investor bargained for); Elizabeth Snodgrass, *Protecting Investors’ Legitimate Expectations: Recognizing and Delimiting a General Principle*, 21 ICSID REV.-FOR. INV. L.J. 1, 35, 56–57 (2006); Yannaca-Small, *supra* note 237 at 126.

might rule in Eli Lilly's case and also to underscore the problems with the broader standard of a stable legal and business environment.

1. *Eli Lilly Has No Legitimate Expectation that Common Law Interpretations Will Not Change*

Eli Lilly should not have a claim for violation of fair and equitable treatment under even the broadest standard—that the state maintain a stable legal and business environment—based on prior cases. A frequently cited definition of what constitutes a stable legal and business environment is:

[T]he host State [must] act in a consistent manner, free from ambiguity and totally transparently in its relations with the foreign investor, so that it may know beforehand any and all rules and regulations that will govern its investments, as well as the goals of the relevant policies and administrative practices or directives, to be able to plan its investment and comply with such regulations.²⁵³

This definition has encountered substantial criticism for being impossible to meet.²⁵⁴ It will nonetheless be discussed as a standard that could be applied to establish that even under this overly broad and criticized standard, Eli Lilly should not have a valid claim for violation of fair and equitable treatment.²⁵⁵ Although it is true that the common definition seems to prevent a state from ever changing any laws that may impact a foreign investor without adequate notice, tribunals citing this definition seem to apply it more narrowly.

Specific cases help to put this in context. For example, in *Tecmed v. Mexico*, the tribunal found a violation of fair and equitable treatment because Mexican authorities failed to renew a necessary landfill permit they had

253. *Tecmed Award*, *supra* note 202, ¶ 154; *Metalclad Award*, *supra* note 177, ¶ 99; *MTD Equity Sdn Bhd. v. Republic of Chile*, ICSID Case No. ARB/01/7, Decision on Annulment, ¶ 107 (Mar. 21, 2007); *Siemens A.G. v. Arg. Republic*, ICSID Case No. ARB/02/8, Award, ¶ 297 (Feb. 6, 2007); *GAMI Final Award*, *supra* note 244, ¶ 91; *Occidental Exploration & Prod. Co. v. Republic of Ecuador*, LCIA Case No. UN3467, Final Award, ¶ 185 (BIT/UNCITRAL Arb. Trib. July 1, 2004); *see also* UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at 64 (referring to quoted provision from *Tecmed* as the “classic statement” of this investor-friendly approach).

254. *E.g.*, Zachary Douglas, *Nothing if Not Critical for Investment Treaty Arbitration: Occidental, Eureka and Methanex*, 22 ARB. INT'L 27, 28 (2006) (stating that this is “actually not a standard at all,” but, rather, “a description of perfect public regulation in a perfect world, to which all states should aspire but very few (if any) will ever attain); *White Indus. Austl. Ltd. v. Republic of India*, Final Award, ¶ 10.3.5 (BIT/UNCITRAL Arb. Trib. Nov. 30, 2011) (noting that this statement has been subject to “valid criticism”); *MTD Equity Sdn Bhd. v. Republic of Chile*, ICSID Case No. ARB/01/7, Decision on Annulment, ¶¶ 66–78 (Mar. 21, 2007) (suggesting that this standard is questionable).

255. This is especially true given that Eli Lilly is suing under NAFTA, which links this standard to customary standards of international law.

previously promised to renew.²⁵⁶ In both *CMS v. Argentina* as well as *Enron v. Argentina*, the tribunal found a violation based on this standard where Argentina dismantled its prior regime of tariff guarantees, without which foreign companies would not have made investments in Argentina.²⁵⁷ In *Occidental v. Ecuador*, the tribunal found a violation of fair and equitable treatment based on Ecuador's "manifestly wrong" interpretation of a contract with the investor, as well as Ecuador's inconsistent and unclear value-added tax laws, which negatively impacted the investor, such that the business and legal framework were disrupted.²⁵⁸ In *PSEG v. Turkey*, the tribunal found that Turkey violated the fair and equitable treatment standard because Turkey engaged in inconsistent administrative acts that included ignoring legal rights, as well as a "roller coaster" of continuing legislative changes that negatively affected the investor's power plant.²⁵⁹ The tribunal found that these changes were the exact opposite of stability since the law as well as its interpretation and implementation were continuously changing.²⁶⁰

Although Eli Lilly claims that it was "entitled to rely on the stability, predictability and consistency of Canada's legal and business framework existing at each stage of the establishment, expansion, and development of Lilly's investment" in its drugs,²⁶¹ its claim is far different from prior situations where tribunals found a violation of fair and equitable treatment. Eli Lilly's complaint is unlike the prior decisions in which tribunals found that domestic law specifically induced an investor to make investments in the country that were then negatively impacted by a change in law. Indeed, Eli Lilly has made no allegation that Canada's prior law induced Eli Lilly to make any investments. In addition, Canada's proper application of current law on utility requiring the promise doctrine to be met is neither a "manifestly wrong" legal interpretation nor a "roller coaster" of changes. Canadian courts have not engaged in any manifestly wrong legal interpretations; to the contrary, courts have consistently and correctly ruled against Eli Lilly based on prevailing law. In addition, while this doctrine is arguably different than

256. Tecmed Award, *supra* note 202, ¶ 165–66.

257. *CMS Gas Transmission Co. v. Arg. Republic*, ICSID Case No. ARB/01/8, Award (May 12, 2005) [hereinafter *CMS Gas Award*]; *Enron Corp. v. Arg. Republic*, ICSID Case No. ARB/01/3, Award (May 22, 2007) [hereinafter *Enron Award*]. However, it may have been relevant in these cases that the tribunal was applying an agreement that specifically noted that the standard is "desirable in order to maintain a stable framework for investment." *CMS Gas Award*, *supra*, ¶ 274; *Enron Award*, *supra*, ¶¶ 259–60.

258. *Occidental*, *supra* note 253, ¶ 184.

259. *PSEG Global Inc. v. Republic of Turk.*, ICSID Case No. ARB/02/5, Award, ¶¶ 246–50 (Jan. 19, 2007).

260. *Id.* ¶ 254.

261. *Eli Lilly Notice of Arbitration*, *supra* note 2, ¶ 82.

when Eli Lilly first applied for its patents, one modification to the common law definition of utility is a far cry from the multitude of changes considered a problem in *PSEG*. Moreover, in past cases where tribunals have applied this broad standard, the fact that the relevant agreements specifically referenced stability as a goal of the treaty may have influenced the tribunals.²⁶²

2. *An Issued Patent Is Not a State Representation of Permanent Validity That Can Be Justifiably Relied On and Must Be Balanced Against State Interests*

A number of tribunals reject the broad standard of a stable legal and business framework as unrealistic²⁶³ and unfair,²⁶⁴ and instead only recognize claims based on legitimate investor expectations if those expectations outweigh state interests. As the tribunal in *Saluka v. Czech Republic* stated, “[n]o investor may reasonably expect that the circumstances prevailing at the time the investment is made remain totally unchanged. . . . [T]he host State’s legitimate right subsequently to regulate domestic matters in the public interest must be taken into consideration as well.”²⁶⁵ Under such an approach, tribunals only find a claim if it arises from (a) a state’s specific representations or commitments to an investor which the investor has relied on, and only after (b) the investor’s expectations are balanced against legitimate regulatory activities of host countries.²⁶⁶ As this section will explain, Eli Lilly has no valid claim because there was no specific representation by Canada that Eli Lilly was justified in relying on, and Canada had legitimate interests in modifying its patent laws.

262. *E.g.*, *Occidental*, *supra* note 253, ¶ 183 (referring to the preamble); *CMS Gas Award*, *supra* note 257, ¶ 274 (referring to the preamble). In addition, the Argentine cases involved licenses granted by the government which stated that they could not be modified without the licensee’s consent. These licenses may have also played a role in the tribunal’s finding of a breach of fair and equitable treatment.

263. *See, e.g.*, *Douglas*, *supra* note 254, at 28; *El Paso Energy Int’l Co. v. Arg. Republic*, ICSID Case No. ARB/03/15, Award, ¶¶ 352, 371 (Oct. 31, 2011) [hereinafter *El Paso Energy Award*] (noting that “[e]conomic and legal life is by nature evolutionary” such that it is important to consider whether changes to a legal framework are unreasonable or contrary to a “specific commitment”).

264. UNCTAD, *FAIR AND EQUITABLE TREATMENT*, *supra* note 237, at 67. Indeed, one tribunal stated it would be unconscionable for a country to promise not to change its legislation as time and needs change. Even where the agreement’s preamble noted the importance of the stability of a legal framework, the tribunal declined to apply this standard. *Cont’l Cas. Co. Award*, *supra* note 209, ¶ 258.

265. *Saluka Partial Award*, *supra* note 209, ¶¶ 304–08.

266. *E.g.*, *Duke Energy Electroquil Partners v. Republic of Ecuador*, ICSID Case No. ARB/04/19, Award, ¶ 340 (Aug. 18, 2008); *Cont’l Cas. Co. Award*, *supra* note 209, ¶ 261.

a) A Patent Is Not a State Representation of Guaranteed Validity

The first question is what constitutes a state representation. The most typical state representation is a specific state commitment to an investor. As with expropriation claims, a tribunal may find a state commitment exists if there is some action attributable to the state, such as a representation from a government official.²⁶⁷ In addition, this state action must be either a specific commitment to the investor, or a set of general rules put in place with the intention of inducing foreign investment upon which the investor relied.²⁶⁸ In this case, there is no suggestion that prior Canadian law was intended to induce foreign investment, such that inducement is not discussed.²⁶⁹

The issue here is whether Canada made a specific commitment to Eli Lilly. Tribunals and scholars consider a commitment specific if its “precise object was to give a real guarantee of stability to the investor.”²⁷⁰ Accordingly, general statements in treaties or legislation do not suffice.²⁷¹ On the other hand, a specific commitment could include a commitment made in a contract or letter,²⁷² or an explicit promise or guarantee from the state.²⁷³ However, although there are heightened expectations from a contract, not every breach of a contract necessarily violates the fair and equitable treatment standard;²⁷⁴ rather, something more is necessary, although tribunals have not

267. Metalclad Award, *supra* note 177, ¶ 73 (noting that it was unchallenged that state and local government acts were attributable to the state); *see also* Stephen Fietta, *Expropriation and the “Fair and Equitable” Standard: The Developing Role of Investors’ “Expectations” in International Investment Arbitration*, 23 J. INT’L ARB. 375 (2006).

268. *E.g.*, Glamis Gold, Ltd. v. United States, Award, ¶ 627 (NAFTA/UNCITRAL Arb. Trib. June 8, 2009) [hereinafter *Glamis Gold*].

269. *E.g.*, Enron Award, *supra* note 257, ¶¶ 264–67; LG&E Energy Corp. v. Arg. Republic, ICSID Case No. ARB/02/1, Award, ¶¶ 132–39 (July 25, 2007).

270. El Paso Energy Award, *supra* note 262, ¶ 377 (emphasis removed).

271. *E.g.*, Cont’l Cas. Co. Award, *supra* note 209, ¶ 261; *see also* El Paso Energy Award, *supra* note 259, ¶ 394 (noting that this would “immobilize the legal order and prevent any adaptation to circumstances”).

272. *E.g.*, El Paso Energy Award, *supra* note 262, ¶ 376. However, breach of a contract is not per se a violation of a specific commitment. UNCTAD, FAIR AND EQUITABLE TREATMENT, *supra* note 237, at 87.

273. Parkerings Final Award, *supra* note 232, ¶ 331.

274. *E.g.*, Duke Energy Electroquil Partners v. Republic of Ecuador, ICSID Case No. ARB/04/19, Award, ¶ 358 (Aug. 18, 2008); Gustav F W Hamester GmbH & Co KG v. Republic of Ghana, ICSID Case No. ARB/07/24, Award, ¶ 335 (June 18, 2008) [hereinafter *Hamester Award*]; Impregilo S.p.A. v. Arg. Republic, ICSID Case ARB/07/17, Final Award, ¶ 181 (June 21, 2011); *see also* Parkerings Final Award, *supra* note 232, ¶ 344 (“The expectation a party to an agreement may have of the regular fulfillment of the obligation by the other party is not necessarily an expectation protected by international law. . . . [T]he party whose contractual expectations are frustrated should . . . seek redress before a national tribunal.”).

necessarily been consistent in assessing what additional activity suffices.²⁷⁵ One tribunal suggested that there must be a denial of justice or discrimination.²⁷⁶ Another tribunal found a violation of this standard not just based on a simple breach, but the fact that the state took action inconsistent with the investment agreement for an urban development project by denying relevant permits to complete the project.²⁷⁷

Importantly, a mere expectation that the law will not change does not constitute a specific commitment made by the state that lays the groundwork for a violation of fair and equitable treatment.²⁷⁸ For example, in *Methanex*, the tribunal held there was no violation of the standard of fair and equitable treatment standard when California changed its laws to ban certain carcinogenic additives to methanol and that change essentially destroyed the investor's market because there was no representation that regulatory laws would not change.²⁷⁹ Similarly, a tribunal found that Canadian company Glamis had no legitimate expectation that the United States (through California) would not pass legislation that would impact Glamis's mining investment, even when California's action was a significant change from settled practice, where California made no specific statements to induce investment.²⁸⁰ Also, in *ADF v. United States*, the tribunal found the investor had no legitimate expectation that the law would remain unchanged when the state made no representation and the investor instead simply relied on advice by private counsel.²⁸¹ More recently, in *Total v. Argentina*, a tribunal held that the legal regime in force at the time an investment is made is not guaranteed to remain in force unless the state has explicitly assumed a legal obligation, such as a stabilization clause.²⁸² Tribunals have noted that absent unusual situations, such as a drastic or discriminatory change in laws, there should be

275. *E.g.*, Potestà, *supra* note 252, at 15–18.

276. Glamis Gold, *supra* note 268, ¶ 620.

277. MTI Equity Sdn Bhd. v. Republic of Chile, ICSID Case No. ARB/01/7, Decision on Annulment (Mar. 21, 2007).

278. El Paso Energy Award, *supra* note 262, ¶ 371.

279. Methanex Final Award, *supra* note 209, at Section IV.D, ¶ 7.

280. Glamis Gold, *supra* note 268, ¶¶ 766–67, 801–02; *see also* Parkerings Final Award, *supra* note 332, ¶ 334–38 (finding no violation of legitimate expectations that Lithuania would not change its laws given that Lithuania was transitioning from being part of the Soviet Union to becoming a candidate for EU membership).

281. ADF Grp. Inc. v. United States, ICSID Case No. ARB(AF)/00/1, Final Award (Jan. 9, 2003).

282. Total S.A. v. Arg. Republic, ICSID Case No. ARB/04/1, Decision on Liability, ¶¶ 117, 429 (Dec. 27, 2010).

no liability under the fair and equitable treatment standard when there is no stabilization clause.²⁸³

A key question with respect to Eli Lilly's claims is thus whether Canada made any specific representations to Eli Lilly that Eli Lilly relied on. The only possible representation stems from Eli Lilly's novel claim that the issued patents are a contract, such that the patent itself is a representation that the patent will never be revoked.²⁸⁴ However, unlike a contract, which can generally be canceled only in extreme circumstances, issued patents are only presumptively valid and are often canceled if found to fail to meet one of the required criteria.²⁸⁵ Moreover, as noted earlier, even a breached contract with a state is not necessarily enough to establish a violation of the fair and equitable treatment standard. Something more is usually necessary.²⁸⁶

b) There Has Been No Negative Reliance upon a State Representation

Even if there is a state representation, it is important that there be reliance on that representation to the investor's detriment due to induced investments.²⁸⁷ For example, in *Metalclad*, the investor relied on the representation of officials that the investor had all necessary federal and state permits to construct a hazardous waste landfill and expended capital in constructing the landfill.²⁸⁸ Thus the denial of the municipal construction permit violated the investor's legitimate expectations.²⁸⁹

Eli Lilly has no viable argument that it relied on commitments that induced it to invest in developing its invention and to apply for a Canadian patent. Eli Lilly seems to complain that it could not have expected Canada to modify domestic standards when it applied for a patent. However, there was no specific representation that Canadian law would not change when Eli Lilly applied. Moreover, Eli Lilly's expenditure of capital to develop the drug it sought to patent is not tied to Canadian laws. As mentioned earlier, multinational pharmaceutical companies develop drugs that they aim to

283. *E.g.*, *Toto Costruzioni Generali S.p.A. v. Republic of Leb.*, ICSID Case. No. ARB/07/12, Award, ¶ 244 (June 7, 2012).

284. *Eli Lilly Notice of Arbitration*, *supra* note 2, ¶ 82.

285. *Cf.* NAFTA, *supra* note 36, art. 1110(7) (expressly recognizing invalidated patent claims as outside the realm of expropriation, which suggests that there are no legitimate expectations that a patent will never be invalidated).

286. *E.g.*, *Parkerings Final Award*, *supra* note 232, ¶ 344; *Hamester Award*, *supra* note 270, ¶ 337.

287. *E.g.*, *El Paso Energy Award*, *supra* note 262, ¶ 376.

288. *Metalclad Award*, *supra* note 177, ¶¶ 85–88.

289. *Id.* ¶¶ 89–90, 99–101.

patent in any and all countries that will provide such patents. In addition, even if Eli Lilly claimed that it was induced to invest in promoting its new drug, this claim should also fail because an issued patent is not a guarantee that the patent will remain valid. Given that it is common for issued patents to be invalidated, the existence of a patent should not induce investment in promoting a drug.

c) Eli Lilly Has No Legitimate Expectation That Outweighs
Canada's Interests

The final method of determining legitimate expectations requires balancing investor expectations against state policy. The facts of some past tribunal cases may help to shed light on how this balance applies. For example, although the tribunal in *Saluka* recognized the importance of considering legitimate regulatory action, the tribunal found that the Czech Republic had no legitimate reason to protect similarly situated domestic, but not foreign, banks.²⁹⁰ In contrast, in *EDF v. Romania*, the tribunal found that a statute passed to abolish duty-free operations in Romanian airports was a reasonable response to the legitimate problem of contraband and did not disproportionately or discriminatorily impact the claimant's investments since it applied equally to all operators.²⁹¹ In addition, some tribunals suggest that there should be a high level of deference to states to regulate matters within their own borders.²⁹²

Past decisions favor a finding that Canada has a legitimate interest in the current promise doctrine. First, Canada does have a bona fide interest in promoting fundamental patent policy that ensures patents are only issued when there is adequate disclosure to justify the social cost of a patent.²⁹³ In addition, Canada's law applies equally to all foreign and domestic companies. Even though all pharmaceutical companies are implicated, the impact on this single industry is no different than the situation in *Saluka* where all owners of duty-free operations were impacted. In *Saluka*, the tribunal found that Romania responded reasonably to a contraband problem by enacting a law that impacted all owners of duty free operations. Here, Canadian courts appropriately responded to the problem of how to ensure that Canadian

290. *Saluka* Partial Award, *supra* note 209, ¶¶ 304–08.

291. *Id.* ¶¶ 293–94; *see also* *EDF (Servs.) Ltd.*, *supra* note 245, ¶ 219 (noting that “[l]egitimate expectations cannot be solely the subjective expectations of the investor” and that proper consideration of “the host State’s power to regulate its economic life in the public interest” should be taken into account).

292. *S.D. Myers* Partial Award, *supra* note 135, ¶ 263; *GAMI* Final Award, *supra* note 244, ¶ 93.

293. *See supra* Section II.A (explaining the importance of patent disclosures).

patents serve the traditional purpose of requiring proper disclosure of an invention before burdening the public with a patent by creating a doctrine that arguably impacts pharmaceutical innovations more. Although the expropriation doctrine provides some protection to foreign investors, it is not intended to provide better protection than that afforded to domestic investors.

III. BEYOND ELI LILLY'S CASE: PENDING PROBLEMS AND HOW TO ADDRESS THEM

This Part goes beyond the *Eli Lilly* case to highlight other domestic laws at the intersection of intellectual property and public health vulnerable to challenge in investor-state arbitration proceedings. In particular, this Part explains TRIPS-consistent domestic actions that might nonetheless result in investor-state claims. After explaining claims that companies (which qualify as foreign investors) are likely to make in investor-state proceedings, this Part provides specific proposals that can be incorporated in pending agreements to minimize these problems.

A. PUBLIC HEALTH ISSUES IN DANGER OF DISRUPTION

Eli Lilly's suit may portend the beginning of a trend towards investor challenges to a number of controversial issues concerning the balance of pharmaceutical interests and public health. These issues include patentability criteria beyond the one challenged in *Eli Lilly*, issuance of compulsory licenses on patents, and domestic regulations concerning protection of clinical data submitted to obtain approval to sell drugs.

1. Patentability Standards and Compulsory Licenses

One patent standard that is especially vulnerable to challenge under investor-state arbitration is a criteria that makes new drugs that are very similar to existing drugs unpatentable unless they show improved efficacy.²⁹⁴ Companies and lawyers alike have improperly suggested that section 3(d) of India's patent law barring patents on drugs that are very similar to existing drugs without providing increased efficacy is inconsistent with TRIPS.²⁹⁵ In

294. *E.g.*, The Patents (Amendment) Act, 2005, No. 15, § 3(d), Acts of Parliament, 2005 (India) (clarifying that inventions do not include "the mere discovery of a new form of a known substance which does not result in the enhancement of the known efficacy of that substance or the mere discovery of any new property or new use for a known substance or of the mere use of a known process, machine or apparatus unless such known process results in a new product or employs at least one new reactant").

295. *E.g.*, JONES DAY, *supra* note 19, at 19; *see also* PHRMA, SPECIAL 301 SUBMISSION 2014, at 26 (referring to 3(d) as an "additional hurdle"); K. M. Gopakumar, *Intellectual Property*

the eight years since India pioneered this law there have been no challenges to its TRIPS consistency in the WTO forum. However, while countries tend to be hesitant to bring WTO disputes due to political considerations and concern with possibly undesirable precedent, companies do not share these concerns in seeking investment remedies.²⁹⁶ Accordingly, India's and other similar laws are ripe for challenge by foreign companies to the extent that there is an applicable investment agreement. Even in the absence of a specific challenge, Eli Lilly's suit alone could make a country hesitant to adopt such laws given the potential cost of a challenge,²⁹⁷ in addition to potential awards granted as compensation to foreign investors who establish violation of investment claims.²⁹⁸

A company could claim that India's section 3(d) patent standard results in improper expropriation, including a claim that the law is not consistent with TRIPS in a manner similar to Eli Lilly's case. As noted earlier, most agreements technically exclude from the definition of expropriation any domestic denial of intellectual property rights if the denial is consistent with TRIPS. However, some have suggested India's section 3(d) provision imposes an additional patentability requirement not permitted by TRIPS, and thus violates TRIPS.²⁹⁹ This is incorrect given that India has not imposed a

Issues Dominate the USITC Public Hearing on India, THIRD WORLD NETWORK (Feb. 13, 2014), <http://www.twinside.org.sg/title2/health.info/2014/hi140201.htm> (noting PhRMA claim that 3d adds additional requirements).

296. *E.g.*, Ermias Tekeste Biadgleng, IP Rights under Investment Agreements: The TRIPS-Plus Implications for Enforcement and Protection of Public Interest, South Centre Research Paper No. 8, at 30 (2006), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=943013. Moreover, law firms seem to encourage such action. *E.g.*, JONES DAY, *supra* note 19; Nandan S. Nelivigi et al., *Indian Patent Office Grants Compulsory License for Bayer's Nexavar: Implications for Multinational Drug Companies*, WHITE & CASE (Apr. 2012), <http://www.whitecase.com/alerts-04022012/>.

297. *See, e.g.*, SCOTT SINCLAIR, CAN. CTR. FOR POLICY ALTS., NAFTA CHAPTER 11 INVESTOR-STATE DISPUTES 24–25 (2010), available at <https://www.policyalternatives.ca/sites/default/files/uploads/publications/National%20Office/2010/11/NAFTA%20Dispute%20Table.pdf> (noting that whereas the cost of simply administering a panel can be \$ 1 million or more, legal costs can be substantially higher and even with respect to frivolous claims that never get a full hearing can still cost several hundred thousand dollars).

298. For example, Mexico was ordered to pay compensation of nearly \$170 million (plus interest) in 2009 for three decisions involving a Mexican tax on high fructose corn syrup. This exceeded the total annual GDP of the poorest sixteen Mexican states. SINCLAIR, *supra* note 297. In 2010, the Philippines spent nearly \$60 million to defend two cases against a German investor. Kim Arveen M. Patria, *Study Investment Provisions Before an FTA Says Advocacy Group*, FOCUS ON THE GLOBAL SOUTH, <http://focusweb.org/content/study-investment-provisions-fta-says-advocacy-group> (last visited Dec. 15, 2014).

299. *E.g.*, *supra* note 295 and accompanying text; *see also In re Natco Pharma. Ltd. & Bayer Corp.*, C.L.A. No. 1 of 2011 (Controller of Patents Mar. 9, 2012) (India); Bayer Corp.

new requirement of patentability; rather, India has simply provided a different definition of patentability that is permissible and in fact contemplated by TRIPS. Just as TRIPS permits Canada to define what is “useful” for Canada’s patents laws, TRIPS permits India to define what is an “invention,” as well as what is “new,” such that a number of scholars and policymakers consider India’s laws to be consistent with TRIPS.³⁰⁰ Nonetheless, just as Eli Lilly has incorrectly challenged Canada as violating NAFTA with an investment agreement, companies are similarly likely to challenge India’s TRIPS-consistent standard.

A possibly even bigger problem is that countries that want to copy India’s law may face claims by companies that the companies have been denied fair and equitable treatment due to an undesirable change in the law. An unduly broad interpretation of such claims might permit an investor to recover if a country changed its laws in a way that altered the legal environment. As noted earlier, there should not be any legitimate expectation that the law will never change. Nonetheless, companies win the vast majority of these cases, such that any potential claim—including an unsubstantiated one—could chill pending proposals for reform of patent laws.

2. *Compulsory Licenses*

Another likely target for investor-state arbitration is a compulsory license. A compulsory license is a traditionally recognized, state-mandated license to use a patented invention in certain instances; the patent is still valid, but the patent owner cannot exclude the licensee and must accept the government dictated royalty.³⁰¹ Although this situation seems inapposite of the patent right to exclude, one of the reasons compulsory licenses have historically been granted is to promote public interest on a number of grounds, including access to medicine.³⁰² The ability to issue compulsory licenses is especially important now to developing countries with limited resources; previously, they could promote access to low-cost drugs by declining to issue patents,

v. Natco Pharma. Ltd., Order No. 45/2013, ¶ 54 (Intellectual Property Appellate Board, Chennai Mar. 4, 2013).

300. *E.g.*, CYNTHIA M. HO, ACCESS TO MEDICINE IN THE GLOBAL ECONOMY: INTERNATIONAL AGREEMENTS ON PATENTS AND RELATED RIGHTS 95–96 (2011); Amy Kapczynski, *Harmonization and its Discontents: Case Study of TRIPS Implementation in India’s Pharmaceutical Sector*, 97 CALIF. L. REV. 1571, 1597–98 (2009).

301. *E.g.*, JEROME H. REICHMAN & CATHERINE HASENZAH, UNITED NATION CONFERENCE ON TRADE & DEV. & INT’L CENTRE FOR TRADE & SUSTAINABLE DEV., NON-VOLUNTARY LICENSING OF PATENTED INVENTIONS 6 (2003).

302. *E.g.*, *id.* at 11–12.

but now that they must issue some patents under TRIPS, compulsory licenses are one of the few tools available to promote lower-cost drugs.³⁰³

Although compulsory licenses are permissible under TRIPS, they are likely to be challenged by patent owners as expropriation. Notably, public statements by pharmaceutical companies often talk about compulsory licenses as either “breaking” their patents, or even expropriating their patent rights.³⁰⁴ Scholars have been expecting such claims.³⁰⁵ This makes sense because a compulsory license may be a prototypical situation where an investor believes that it needs and deserves the additional protection of investor-state arbitration because the investor considers the TRIPS requirements, as well as domestic laws implementing those requirements, inadequate.

Although compulsory licenses that are consistent with TRIPS should technically be exempt from indirect expropriation claims in most cases,³⁰⁶ an

303. See TRIPS, *supra* note 3, art. 27(1). Although nations can impose heightened patentability requirements, it is highly unlikely that this will result in the pre-TRIPS situation of zero patents on any drugs. Indeed, although India has been at the forefront of imposing new restrictions on patentability, the vast majority of pharmaceutical patents have in fact been granted. E.g., C.H. Unnikrishnan, *Foreign Drug Makers Won 77% of All Patents in Last 3 Years*, LIVE MINT (Sept. 15, 2014), <http://www.livemint.com/Industry/H0ZID9XSP78ESv1MAUuZZN/Foreign-drug-makers-won-77-of-all-patents-granted-in-last-3.html>; Broo Baker, *Pharmaceutical Patents Granted by Indian Patent Office (IPO)*, TECH CORP LEGAL LLP, http://techcorplegal.com/Indian_Law_Firm/foreign-pharmaceutical-companies-granted-pharmaceutical-patents-by-indian-patent-office-ipo/ (indicating that over two thirds of pharmaceutical patent applications were granted). The most up to date list of granted and pending applications since July 2012, is available from the Indian patent office at <http://164.100.176.38/tk/> (select “pharmaceutical” from dropdown menu).

304. E.g., Cynthia M. Ho, *Unveiling Competing Patent Perspectives*, 46 HOUS. L. REV. 1047, 1069–70 (2009); see also Merck & Co., Inc. *Statement on Brazilian Government’s Decision to Issue Compulsory License for STOCRIN*, BUSINESS WIRE (May 4, 2007, 12:11 PM), <http://www.businesswire.com/news/home/20070504005566/en/Merck-Statement-Brazilian-Governments-Decision-Issue-Compulsory/> (last visited Aug. 12, 2014) (referring to compulsory license as expropriation); see also JONES DAY, *supra* note 18 (referring to compulsory license as expropriation).

305. E.g., Peter B. Rutledge, *TRIPS and BITs: An Essay on Compulsory Licenses, Expropriation, and International Arbitration*, 13 N.C. J. L. & TECH. ONLINE 149, 161 (2012), available at http://digitalcommons.law.uga.edu/fac_artchop/827/; Gibson, *supra* note 205, at 359.

306. See, e.g., NAFTA, *supra* note 36, art. 1110(7) (excluding compulsory licenses granted consistent with NAFTA, which has provisions similar to TRIPS); U.S.-Singapore Free Trade Agreement, *supra* note 179, art. 15.6.5 (excluding compulsory licenses granted consistent with TRIPS); Dominican Republic–Central America–U.S. Free Trade Agreement art. 10.7.5, Aug. 5, 2004, 43 I.L.M. 514 (2004) [hereinafter CAFTA] (excluding compulsory licenses consistent with TRIPS); see also HOWARD MANN ET AL., IISD MODEL INTERNATIONAL AGREEMENT ON INVESTMENT FOR SUSTAINABLE DEVELOPMENT art. 8(G) (2005)

investor-state dispute would likely still be initiated to assess whether the license in fact complies with TRIPS because TRIPS requirements for compulsory licenses are highly controversial and contested. Part of the problem is that determining whether the licenses comply with TRIPS requires an interpretation of undefined, yet key, terms. For example, TRIPS permits compulsory licenses when a state provides “adequate remuneration,” without any definition of what would be “adequate.” In addition, although TRIPS provides many procedural requirements a state must follow to issue an appropriate license, there is some controversy concerning the ground for issuing a license in the first instance.³⁰⁷

Companies are likely to challenge royalty rates of compulsory licenses as not TRIPS compliant because TRIPS does not clearly define what compensation is “adequate” and companies believe that any compulsory license fails to provide adequate compensation. This is aptly illustrated in the recent case concerning India’s compulsory license on Bayer’s cancer drug sold as Nexavar. Bayer sought a royalty rate of fifteen percent of net sales whereas the court granted a royalty of six percent. Although a subsequent appeal raised the royalty to seven percent, that rate is still less than half of what the patent owner sought.³⁰⁸ Even though Bayer strongly contested the royalty rate, that rate was completely within the guidelines issued by the World Health Organization and the United Nations Development Programme.³⁰⁹

One law firm suggested that the six percent royalty rate constituted indirect expropriation that failed to provide adequate compensation to Bayer pursuant to a typical investment agreement that requires compensation “equivalent to the value of the expropriated . . . investment immediately before the date on which such expropriation . . . became publicly known.”³¹⁰ Although some might suggest that there is no conflict between this expropriation standard and the ambiguous TRIPS requirement of “adequate”

(excluding from expropriation compulsory licenses granted consistent with “applicable international agreements on intellectual property”).

307. TRIPS, *supra* note 3, art. 31.

308. *In re Natco Pharma. Ltd. & Bayer Corp.*, C.L.A. No. 1 of 2011 (Controller of Patents Mar. 9, 2012) (India); *Bayer Corp. v. Natco Pharma. Ltd.*, Order No. 45/2013, ¶ 54 (Intellectual Property Appellate Board, Chennai Mar. 4, 2013).

309. See JAMES LOVE, WORLD HEALTH ORGANIZATION, REMUNERATION GUIDELINES FOR NON-VOLUNTARY USE OF A PATENT ON MEDICAL TECHNOLOGIES (Robert Weissman ed., 2005).

310. JONES DAY, *supra* note 19, at 3.

compensation, scholars generally believe that “adequate” compensation under TRIPS is not intended to be market rate.³¹¹

Another TRIPS-consistent aspect of compulsory licenses that foreign investors could challenge is the grounds for issuing a compulsory license in the first instance. TRIPS permits countries to decide the basis for issuing compulsory licenses and only governs procedural aspects of compulsory licensing. However, there have been many misstatements concerning permissible grounds for issuing compulsory licenses made not only by companies, but also by scholars and government officials.³¹² Contrary to the common claim by companies that compulsory licenses are only appropriate in case of an emergency or if the patent owner cannot provide adequate supply of the needed drug,³¹³ countries have complete discretion to decide the grounds for issuing compulsory licenses.³¹⁴ This is very important for countries like India that have unusual legal criteria for issuing a compulsory license, such as a drug not being available at a “reasonably affordable price” from the patent owner.³¹⁵

Notably, even if an arbitration tribunal were to properly find that the above two issues were consistent with TRIPS, and thus not indirect expropriation, a tribunal might still find a violation of fair and equitable treatment. There is no intellectual property exception to fair and equitable treatment claims for even TRIPS-consistent measures in any existing agreement. A company might argue that it applied for a patent to its detriment because it did not expect that a country would issue a compulsory license that demolished the value of its patent. A tribunal that took a broad

311. *E.g.*, HO, *supra* note 300, at 138; Tsai-Yu Lin, *Compulsory Licenses for Access to Medicines, Expropriation and Investor-State Arbitration Under Bilateral Investment Agreements—Are There Issues Beyond the TRIPS Agreement?*, 40 IIC: INT'L REV. INTEL. PROP. & COMPETITION L. 152, 163–64 (2009); Antony Taubman, *Rethinking TRIPS: 'Adequate Remuneration' for Non-Voluntary Patent Licensing*, 11 J. INT'L ECON. L. 927, 951–55, 957 (2008) (explaining that this is not necessarily equivalent to full market value); Biadgleng, *supra* note 296, at 18; *see also* Gibson, *supra* note 205, at 415 (suggesting that it is unclear whether a TRIPS consistent license would constitute expropriation under an investment agreement).

312. *E.g.*, Cynthia M. Ho, *Patent Breaking or Balancing?: Separating Strands of Fact From Fiction under TRIPS*, 34 N.C. J. INT'L L. & COM. REG. 371 (2009); *see also* James Love, USPTO “Clarifies” June 27, 2012 Testimony on Biologics Exclusivity and India Compulsory License, KNOWLEDGE ECOLOGY INT'L, (July 3, 2012), available at <http://www.keionline.org/node/1452>.

313. *E.g.*, JONES DAY, *supra* note 19, at 3.

314. TRIPS, *supra* note 3, art. 31; Doha Public Health Declaration, *supra* note 34, ¶ 5(b) (“Each member has the right to grant compulsory licenses and the freedom to determine the grounds upon which such licenses are granted.”); *see also* HO, *supra* note 300, at 128–29.

315. The Patents (Amendment) Act, 2005, No. 15, § 84, Acts of Parliament, 2005 (India).

view of the fair and equitable treatment standard to demand a stable legal environment might be sympathetic to such a claim. Even if there is no technical change in domestic laws, if a country had simply not previously issued compulsory licenses, or rarely issued such licenses, a company might nonetheless complain that an unexpected issuance of a compulsory license was inconsistent with fair and equitable treatment. Given that such claims are unpredictable and highly successful for claimants, there is a serious risk that tribunals deciding investor-state claims would nonetheless find that a TRIPS-consistent license is an expropriation or a violation of fair and equitable treatment.

3. *Domestic Regulation of Clinical Data*

In addition, nations may be subject to investment claims concerning domestic regulations governing clinical data relating to new drugs. There are two related issues that could be subject to challenge. First, companies may bring challenges against countries that permit generic applicants to immediately rely on clinical data without providing a period of “data exclusivity” for the initial drugmaker. Second, companies may challenge domestic laws requiring that all clinical trials of approved drugs be made publicly available. Although both potential challenges relate to the same data, they will be discussed separately because they involve separate issues (reliance versus disclosure), as well as different issues of interpretation under TRIPS.

a) Countries That Do Not Provide “Data Exclusivity” Will Likely Be Challenged

To best understand the data exclusivity issue, some background concerning the regulatory drug approval process is necessary. Unlike most other patented items, patented drugs need regulatory approval by a domestic agency such as the U.S. Food and Drug Administration before the drugs can be sold. Most countries grant such approval when a company can establish that the proposed new drug is safe and effective for its proposed use based on clinical data.³¹⁶ It can take many years and millions of dollars to compile the requisite data for regulatory review.³¹⁷

In contrast, manufacturers of proposed generics can gain approval with a more limited set of clinical data. Most countries will approve generic versions

316. *E.g.*, Federal Food, Drug, and Cosmetic Act, 21 U.S.C. § 355(b)(1) (2012) (U.S. standard); Food and Drug Regulations, C.R.C., c. 870, § C.08.0002(2) (Can.) (Canadian standard).

317. *E.g.*, Salomeh Keyhani et al., *Are Development Times for Pharmaceuticals Increasing or Decreasing?*, 25 HEALTH AFF. 461, 463 (2006).

based solely on clinical studies that show “bioequivalence” to a previously approved drug; the proposed generic is then presumed to be just as safe and effective as the previously approved drug.³¹⁸ The time and investment needed to establish clinical data of bioequivalence is a mere fraction of the time and investment required to produce data for the earlier drug’s regulatory approval process.³¹⁹ This is an intentional policy decision. After all, a company that is a second or later entrant to the market with no possible patent protection cannot charge high prices to recoup an expensive investment. Moreover, if generic companies are not provided a less costly regulatory approval process, original companies can continue to sell their drugs at premium prices long after a patent has expired due to lack of competition.³²⁰

The issue with data exclusivity is when generic companies can rely on clinical data of the drug they are copying. In a country that provides data exclusivity, the generic manufacturer is barred from relying on that data for a certain period, ranging from five to ten years from approval of the prior drug.³²¹ Data exclusivity, when available, is completely separate from patent protection and can provide substantial commercial advantage for even unpatentable products. In contrast, a country that does not recognize data exclusivity will permit other companies to immediately rely on this data. This means that as soon as a patented drug is approved for sale, a generic manufacturer can apply to sell a lower-cost equivalent. Importantly, this does not mean that the patent is not valid. However, it does permit the manufacturer of a generic to enter the market while simultaneously challenging the patent. Although this may seem like a formidable challenge, the vast majority of challenged drug patents are in fact found invalid or not infringed.³²²

318. *E.g.*, 21 U.S.C. § 355(j)(4)(F) (2012); 21 C.F.R. § 314.92(a)(1) (2010); *Notice: Release of the Health Canada Draft Guidance Document: Quality (Chemistry and Manufacturing): New Drug Submissions (NDSs) and Abbreviated New Drug Submissions (ANDSs)*, HEALTH CANADA (Jan. 24, 2014), http://www.hc-sc.gc.ca/dhp-mps/alt_formats/pdf/consultation/drug-medic/qual_ndsands_draft_pdnpadn_ebauche-eng.pdf.

319. *E.g.*, Henry G. Grabowski et al., *Entry and Competition in Generic Biologics*, 28 *MANAGERIAL & DECISION ECON.* 439, 443 (2007).

320. Indeed, this was the situation in the United States before laws were amended to permit generic drug approvals based on the abbreviated process. *See, e.g.*, Gerald J. Mossinghoff, *Overview of the Hatch-Waxman Act and Its Impact on the Drug Development Process*, 54 *FOOD & DRUG L.J.* 187, 187 (1999).

321. *E.g.*, 21 U.S.C. § 314.108 (2012) (five-year exclusivity); Council Directive 2004/27/EC. Amending Directive 2001/83/EC on the Community Code Relating to Medicinal Products for Human Use, 2004 O.J. (L 136) 34, 39 art. 10.1 (ten-year maximum exclusivity).

322. *E.g.*, EC PHARMACEUTICAL SECTOR INQUIRY, *supra* note 10, para. 501 (finding that in over half of cases where patent was challenged by a generic company, the patent was

A patent-owning company such as Eli Lilly may assert that in a country without data exclusivity, its right to prevent other companies from using its data was indirectly expropriated. Clinical data that is expensive to develop seems to easily fall within the definition of an investment. The expropriation issue is whether permitting generic companies to rely on clinical data results is a substantial and unreasonable interference with this investment. A company would likely believe that this is the case if it cannot completely exclude competitors from using its proprietary data. This seems somewhat analogous to a compulsory license of a patent in that, although the company can still use its data (or patent), lack of exclusivity results in a substantial interference with the expected value. A developing country should have a strong policy argument for denying data exclusivity to permit generic companies to more quickly rely on this data and promote faster entry of low cost drugs. However, it is unclear how a tribunal would balance interests or whether it would instead use the “sole effect” test, which would clearly favor only corporate interests. Accordingly, an important issue is whether a domestic decision to reject data exclusivity could be exempt from consideration as an expropriation claim.

An initial question is whether such a claim could be excluded under clauses that exempt certain intellectual property issues from indirect expropriation claims.³²³ The issue is whether lack of data exclusivity should be considered a “limitation” of “intellectual property rights” pursuant to agreements that exclude such limitations of intellectual property rights from the scope of expropriation. Although data exclusivity is not a traditional intellectual property right, many companies as well as countries consider it to be one in contexts beyond investor-state disputes.³²⁴ However, the phrase

invalidated); FED. TRADE COMM’N, *GENERIC DRUG ENTRY PRIOR TO PATENT EXPIRATION: AN FTC STUDY 20* (2002) (finding that generic companies win in seventy-three percent of challenges, with twenty-eight percent of patents found invalid, thirty-five percent of the cases finding lack of infringement, and ten percent of cases abandoned by the patent owner before a judicial finding); *see also* Wullianallur Raghupathi, *Pharmaceutical Patent Validity: An Empirical Study of the Recent Decisions of the U.S. Court of Appeals for the Federal Circuit (2008-2011)*, 18–19 (Fordham Univ., Working Paper), *available at* http://ghimc.org/pdf/patent_enforcement.pdf (finding that during the period 2005–2006 about half of patents challenged by generics were held invalid).

323. Arguably, another reason the expropriation claim should be excluded is that lack of implementation of a desired law may not constitute state action that is fundamental to an expropriation claim. Generally, expropriation claims are based on an affirmative act, rather than an omission. However, as Eli Lilly’s case shows, companies are not afraid to make new claims in the area of investment arbitrations.

324. Indeed, there are some free trade agreements that require countries to provide data exclusivity under intellectual property chapters. *See, e.g.*, U.S.-Singapore Free Trade

“limitations of intellectual property” suggests that there must be a recognized intellectual property right that can be limited, such that a country that completely denied a right might not fall under this language. Nonetheless, it is at least conceivable that data exclusivity or lack of data exclusivity would be covered as intellectual property.

Even if data exclusivity is considered a type of intellectual property right that falls within the intellectual property exception to expropriation, it is not necessarily immune to challenge. In particular, this exception only applies to intellectual property rights consistent with TRIPS and there is significant controversy concerning what TRIPS requires. In particular, although some companies and countries believe that TRIPS requires data exclusivity, a proper interpretation of TRIPS pursuant to the customary rules of interpretation of international agreements establishes that this view is incorrect.³²⁵ TRIPS requires that countries “protect” data submitted to government for approval of pharmaceuticals from “unfair commercial use” without specifying what this means.³²⁶ Although companies suggest that it is unfair to allow other companies to rely on their data, negotiators rejected language that specifically stated that there could be no reliance on the data.³²⁷ The rejection of this earlier language means that it is not the current standard—contrary to what some companies have suggested.³²⁸ Accordingly, a number of scholars and policymakers consider that the provision does not require data exclusivity.³²⁹

Lack of data exclusivity could also be challenged as a violation of the fair and equitable treatment standard. Although this is recognized as the broadest and most frequently successful claim in investment disputes, it is unlikely to be successful against a country like India. Since India has never recognized

Agreement, *supra* note 179, art. 16.8; Australia-U.S. Free Trade Agreement, U.S.-Austl., art. 17.10, May 18, 2004, 43 I.L.M. 1248 (2004) [hereinafter AUSFTA].

325. *E.g.*, Ho, *supra* note 314, at 76–80.

326. TRIPS, *supra* note 3, art. 39(3).

327. Brussels Draft, ¶4A, reprinted in DANIEL GERVAIS, *THE TRIPS AGREEMENT: DRAFTING HISTORY AND ANALYSIS* 421 (3d ed., 2008); *see also* Bayer Inc. v. Canada (Att’y Gen.) (1998), [1999] 1 F.C. 553 (Can.).

328. *See generally* ORG. OF PHARM. PRODUCERS OF INDIA, *OPPI POSITION PAPER: REGULATORY DATA PROTECTION—A BUILDING BLOCK FOR PHARMACEUTICAL R&D* (2008).

329. *E.g.*, CARLOS CORREA, *TRADE RELATED ASPECTS OF INTELLECTUAL PROPERTY RIGHTS: A COMMENTARY ON THE TRIPS AGREEMENT* 383–92 (2007); JEROME REICHMAN, *UNDISCLOSED CLINICAL TEST DATA UNDER THE TRIPS AGREEMENT AND ITS PROGENY: A BROADER PERSPECTIVE* 10 (2004); Brook K. Baker, *Ending Drug Registration Apartheid: Taming Data Exclusivity and Patent/Registration Linkage*, 34 *AM. J.L. & MED.* 303, 315–16 (2008).

protection for data exclusivity, there would be no legitimate expectation for it to do so even under the broadest standard of maintaining a stable legal and business environment. After all, a stable environment would be the same legal environment.

However, the bigger issue is that the threat of investor-state arbitration could prevent countries from abandoning data exclusivity laws in favor of India's approach, even if a country believes that India's approach is better policy in promoting access to lower-cost drugs. Although tribunals have repeatedly noted that investors should not expect that laws will be frozen in time, a company could claim that they did not expect an existing protection to be dismantled. Some claims could be cabined if tribunals use the more robust standard that only finds violations when an investor relies on a specific state representation since it is unlikely that any country would promise to keep data exclusivity laws. However, that possibility may be too large a risk to take for a developing country with limited funds.

b) Domestic Data Transparency Requirements Are Vulnerable to Challenge

Companies are also likely to challenge domestic regulations concerning disclosure of clinical data supporting approved drugs. The European Union is at the forefront of requiring what is referred to as data "transparency," but if it is challenged, other countries may be hesitant to enact laws that public health scholars uniformly applaud as desirable.³³⁰ In particular, a new EU regulation requires that all clinical data for drugs approved by the European Union be made publicly available.³³¹ Companies strongly oppose disclosing clinical data, claiming that they are entitled to keep such data as a trade secret.³³² Although the regulation is not yet in full effect, companies are likely to contest it once it is.

Before addressing possible claims, it is important to explain the rationale for transparency laws in the context of the regulatory structure for approval of new drugs. As noted earlier, a new drug will be approved for sale based on clinical data that it is safe and effective. Notably, such data is developed not by an independent organization, but by the very company seeking approval.

330. See generally Gardiner Harris, *Diabetes Drug Maker Hid Test Data, Files Indicate*, N.Y. TIMES, July 13, 2010, at A1 (describing problems when drug companies are not required to publicly disclose all relevant data).

331. Regulation 536/2014 of the European Parliament and of the Council on Medicinal Products for Human Use, and Repealing Directive, 2001/20/EC, 2014 O.J. (L 158).

332. Aaron S. Kesselheim & Michelle M. Mello, *Confidentiality Laws and Secrecy in Medical Research: Improving Public Access to Data on Drug Safety*, 26 HEALTH AFF. 483, 483 (2007); see also, e.g., *infra* note 333 (noting AbbView objection).

In addition, although the company must submit the data to the government, the public is not entitled to access it. There are a few cases where independent researchers obtained access to the data either because a country had a policy for doing so in limited circumstances³³³ or because a company responded to public pressure.³³⁴ However, without mandatory transparency, not only doctors and patients, but also governments must rely on industry claims concerning the value of new drugs. Because companies have an interest in selectively publishing positive results,³³⁵ they are more likely to conclude that their drugs are safe and effective than independent researchers.³³⁶ They also overestimate benefits while minimizing risks in published studies.³³⁷ As a result, there may be unnecessary expenditures on expensive new drugs based on questionable data³³⁸ that can also result in negative public health outcomes that could have been avoided.³³⁹ There are a number of examples where new drugs were later found by independent

333. For example, the European Medicines Agency previously granted access to clinical data concerning Abbvie's drug Humira and Intermune's drug Esbriet based on a 2010 policy that such data be provided upon request. European Medicines Agency, European Medicines Agency Policy on Access to Documents (Related to Medicinal Products for Human and Veterinary Use), EMA/110196/2006 (2010). The companies subsequently challenged this decision. Case T-44/13, *AbbVie, Inc. v. European Meds. Agency* (July 17, 2014), <http://curia.europa.eu/juris/document/document.jsf?text=&docid=155902&doclang=en>; Case T-73/13R, *InterMune UK Ltd v. European Meds. Agency* (Apr. 25, 2013), <http://curia.europa.eu/juris/document/document.jsf?text=&docid=137242&doclang=en>; see also Trudo Lemmens, *EMA's Proposed Data Release Policy: Promoting Transparency or Expanding Pharma Control Over Data?*, PLOS BLOGS (May 30, 2014), <http://blogs.plos.org/speakingofmedicine/2014/05/30/emas-new-data-release-policy-promoting-transparency-expanding-pharma-control-data/> (last visited Aug. 12, 2014).

334. E.g., Editorial, *Full Disclosure Needed for Clinical Drug Data*, N.Y. TIMES, July 5, 2013, at A16.

335. E.g., Trudo Lemmens & Candice Telfer, *Access to Information and the Right to Health: The Human Rights Case for Clinical Trials Transparency*, 38 AM. J.L. & MED. 63, 93–94 (2012); Nicholas Bakalar, *Review Finds Drug Makers Issue More Positive Studies*, N.Y. TIMES, Feb. 27, 2007, at F7.

336. E.g., Fujian Song et al., *Dissemination and Publication of Research Findings: An Updated Review of Related Biases*, 14 HEALTH TECH. ASSESSMENT (2010).

337. E.g., Justin E. Bekelman et al., *Scope and Impact of Financial Conflicts of Interest in Biomedical Research*, 289 J. AM. MED. ASSOC. 454 (2003); Joel Lexchin et al., *Pharmaceutical Industry Sponsorship and Research Outcome and Quality: Systematic Review*, 326 BRIT. MED. J. 1167 (2003).

338. Governments stockpiled the antiviral drug to treat influenza sold as Tamiflu based on unverified effectiveness claims by the company that independent researchers only recently determined to be unsubstantiated. E.g., Ben Goldacre, *What the Tamiflu Saga Tells Us About Drug Trials and Big Pharma*, GUARDIAN (Apr. 9, 2014), <http://www.theguardian.com/business/2014/apr/10/tamiflu-saga-drug-trials-big-pharma/>.

339. E.g., HAI EUROPE, *supra* note 31.

research to result in harmful health risks.³⁴⁰ Even though independent researchers can ultimately discover issues, it is expensive, inefficient, and poor public policy to bar them from considering existing data that could result in better public outcomes.

There is a serious concern that transparency requirements would constitute an expropriation. Mandatory disclosure of data would seem to constitute a substantial interference with the expectation that a company's data will not be accessed by a competitor. In addition, the current exceptions of expropriation do not seem to cabin such claims.

A nation should have the right to decide whether to recognize protected data as an intellectual property right. But there is an open issue concerning whether TRIPS requires this data to be protected. In particular, there is a currently untested exception to the TRIPS requirement to protect data from unfair commercial use; TRIPS explicitly states: "Members shall protect such data against disclosure, except where necessary to protect the public, or unless steps are taken to ensure that the data are protected against unfair commercial use."³⁴¹ In other words, TRIPS seems to contemplate that there are in fact some situations where members may not need to protect data against disclosure if necessary to protect the public. Although the European Union may believe that it falls within the TRIPS exception that permits disclosure for public interest, a company would likely believe otherwise.

If a tribunal does not extend the traditional definition of intellectual property to include data exclusivity, such claims could be exempt from a claim for indirect expropriation. This exemption could be based on the language of certain agreements that does not provide a complete exception to expropriation claims, but rather suggests that regulation for public welfare be treated differently. For example, a number of agreements suggest that nondiscriminatory regulatory measures "designed and applied to protect legitimate public welfare objectives, such as public health" do not generally constitute indirect expropriation "except in rare circumstances."³⁴²

In addition, an investor that believes it is entitled to compensation when a country fails to provide data exclusivity raises unique challenges even when public health is involved. Most cases involving public welfare have been

340. *E.g.*, Ho, *supra* note 13, at 501–05.

341. TRIPS, *supra* note 3, art. 39(3).

342. *E.g.*, AUSFTA, *supra* note 324, Annex 11-B, art. 4(b); U.S.-Chile Free Trade Agreement, U.S.-Chile, Annex 10-D, June 6, 2003, 42 I.L.M. 1026; CAFTA, *supra* note 306, Annex 10-C; *Draft, CETA Investment Text, EU-Can., Annex X.11(3): Expropriation*, http://www.tradejustice.ca/wp-content/uploads/2013/08/EU-Canada-FTA-Negotiations-Investment-chapter-4-April-2014_clean.pdf [hereinafter *Draft CETA Investment Text*].

cases that directly impact public health or the environment, such as a regulation that aims to protect sea turtles, or a regulation that aims to reduce carcinogens.³⁴³ In contrast, the public health interest protected in countries that decline to impose data exclusivity is more attenuated. Some consider it obvious that there is not only a universal right to health, but also a right to access to affordable medicine.³⁴⁴ However, there is no universally recognized right to access to affordable medicine; indeed, this is the crux of frequent international tension between companies that want strong patent protection and those that want less patent protection to promote access to affordable drugs.³⁴⁵

343. *E.g.*, *supra* note 213 and accompanying text (discussing how the California ban of an arguable carcinogen should be considered lawful regulation and not expropriation in *Methanex*); KYLA TIENHAARA, THE EXPROPRIATION OF ENVIRONMENTAL GOVERNANCE: PROTECTING FOREIGN INVESTORS AT THE EXPENSE OF PUBLIC POLICY 239–43 (2009); Benjamin W. Jenkins, *The Next Generation of Chilling Uncertainty: Indirect Expropriation Under CAFTA and Its Potential Impact on Environmental Protection*, 12 OCEAN & COASTAL L.J. 269 (2007). Notably, such claims of public welfare do not always succeed in thwarting a claim for expropriation. For example, Mexico was found to have indirectly expropriated Metalclad's investment, and Costa Rica was required to pay \$ 4 million for expropriating land to protect sea turtles. *E.g.*, *Metalclad Award*, *supra* note 177; Joyce Gomez, *Expropriation to Protect Turtles Costs Government \$4 Million*, COSTA RICA NEWS, <http://www.costaricanewssite.com/expropriation-to-protect-turtles-costs-government-4-million/> (last visited Dec. 13, 2014).

344. *E.g.*, Minister of Health v. Treatment Action Campaign 2002 (5) SA 721 (S. Afr.) (right to affordable HIV drugs); Further Initiatives for Social Development, G.A. Res. S-24/2, ¶ 101, U.N. Doc. A/RES/S-24/2 (July 1, 2000); Commission on Human Rights, Res. 2001/33, Access to Medication in the Context of Pandemics Such as HIV/AIDS (Apr. 23, 2001) (asserting right to ARV treatment as part of the right to enjoy the “highest attainable standard of physical and mental health”); High Commissioner of the Comm'n on Human Rights, *The Impact of the Agreement on Trade-Related Aspects of Intellectual Property Rights on Human Rights*, Comm'n on Human Rights, U.N. Doc. E/CN.4/Sub.2/2001/13, ¶¶ 10–15, 27–58 (June 27, 2001); Joanne Csete, *Several for the Price of One: Right to AIDS Treatment as Link to Other Human Rights*, 17 CONN. J. INT'L L. 263, 265 (2002); Lisa Forman, *From TRIPS-Plus to Rights-Plus? Exploring Right to Health Impact Assessment of Trade-Related Intellectual Property Rights Through the Thai Experience*, 7 ASIAN J. WTO & INT'L HEALTH L. & POL'Y 347 (2012); Melissa McClellan, *“Tools for Success”: The TRIPS Agreement and the Human Right to Essential Medicines*, 12 WASH. & LEE J. CIVIL RTS. & SOC. JUST. 153 (2005); Simon Walker, *A Human Rights Approach to the WTO's TRIPS Agreement*, in INTERNATIONAL TRADE AND HUMAN RIGHTS: FOUNDATIONS AND CONCEPTUAL ISSUES 171, 176 (Frederick M. Abbott et al. eds., 2006) (“The right to health . . . includes a right to affordable, essential drugs, as set out in the WHO's essential drug list.”); Alicia Ely Yamin, *Not Just a Tragedy: Access to Medications as a Right Under International Law*, 21 B.U. INT'L L.J. 325 (2003).

345. *E.g.*, Doha Public Health Declaration, *supra* note 34, para. 4 (recognizing that nations should have the ability to promote access to medicine, but without creating a complete exception to patent rights); Amy Kapczynski, *The Access to Knowledge Mobilization and the New Politics of Intellectual Property*, 117 YALE L.J. 804, 852–53 (2008) (describing the use of “frameworks of international rights discourse and corporate malfeasance” for access to medicine, which necessarily suggests that no such universal right currently exists). Although

Countries may face even more problems with a claim for fair and equitable treatment. This standard is often read broadly, and a country that imposes transparency requirements could be considered to be making a substantial change to the legal and business environment. Notably, the EU regulation is not a complete surprise. The European Union has been engaging in increased transparency over the years. As with all such claims, the European Union's interests are better protected under the standard that only recognizes claims based on legitimate expectations due to specific reliance. It is doubtful that the European Union would ever represent that it would not change its laws. However, considering that past cases have broadly interpreted this standard, the EU regulations could still be vulnerable.

B. PROPOSALS TO PRESERVE FLEXIBILITY UNDER TRIPS

This Section provides concrete proposals to address the unique policy issues raised by permitting investors to challenge domestic decisions concerning the proper scope—if any—of intellectual property rights when those decisions are arguably permissible under international agreements such as TRIPS.³⁴⁶ In particular, this Section advocates ideally excluding such issues from international agreements governing investments, or limiting challenges in the dispute settlement system. If this is not possible, this Section suggests specific proposals to cabin expropriation and fair and equitable treatment claims that would otherwise interfere with internationally permissible regulation of intellectual property rights.

the WHO and UN have urged nations to try to expand access to drugs, this remains an issue. See U.N. Secretary General, *Review of the Problem of Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome in All Its Aspects*, ¶ 101, U.N. Doc. A/55/779 (Feb. 16, 2001) (urging states to use compulsory licenses and other means to expand access to lower cost generics). Even those that suggest that corporations be more sensitive to drug prices as a matter of human rights are often cautious to only suggest access for life-saving drugs, and not all drugs. E.g., Chuan-Feng Wu, *Transnational Pharmaceutical Corporations' Legal and Moral Human Rights Responsibilities in Relation to Access to Medicines*, 7 *ASIAN J. WTO & INT'L HEALTH L. & POL'Y* 77 (2012).

346. Although this Section focuses on proposals that stem from the policy issues raised here, there is one issue raised by the Eli Lilly case that impacts all cases: whether procedurally proper decisions of domestic courts should ever be challenged in investor-state disputes. As explained in Part III, there is no precedent or policy reason for enabling investors to obtain compensation using either expropriation or fair and equitable treatment claims when they simply disagree with substantive domestic law. The proposals to limit claims that attempt to challenge substantive law regarding intellectual property equally apply to all other areas.

1. *Exclude Intellectual Property from Investor-State Disputes*

The simplest way to avoid above-noted problems is by narrowing the scope of what constitutes a covered investment. Alternatively, an exception to investor-state disputes could be created to avoid policy problems. In considering these solutions, this Article considers intellectual property to be broadly defined to include not only patents but also any regulatory protection of drugs, such as data exclusivity, since companies themselves consider both to be intellectual property.

There are several approaches to modifying the definition of investment. The most efficient way to eliminate noted problems is to modify the definition of what constitutes an “investment” to explicitly exclude intellectual property rights in their entirety. Not surprisingly, some have suggested doing this.³⁴⁷ It is not adequate for an agreement to simply not mention intellectual property as covered since most agreements cover intangible investments. Even those who oppose Eli Lilly’s claim would need to concede that intellectual property rights are generally considered intangible investments.³⁴⁸ Accordingly, there needs to be an explicit statement that intangible investments do not include intellectual property rights. Alternatively, if intellectual property rights are included as an investment, there should be a clarification that such rights do not include those that have been canceled pursuant to domestic law. Moreover, it may be wise to clarify that domestic law includes common law modifications to the law. This would thus obviate Eli Lilly’s objection that Canada was unjustified in modifying and retroactively applying this standard.

Another possibility is to maintain the scope of covered investments, and instead change the scope of investor-state disputes. In particular, claims that require adjudication of rights under another international agreement, such as TRIPS, could be excluded entirely. Agreements have previously excluded some subject matter, such as national security and tax measures,

347. *E.g.*, Brook K. Baker, *Corporate Power Unbound: Investor-State Arbitration of IP Monopolies on Medicines—Eli Lilly and the TPP*, PJIP RESEARCH PAPER SERIES PAPER 36, at 13 (2013). Although this is obviously a current issue, some countries astutely noted this problem earlier and tried unsuccessfully to eliminate IP from the definition of investments. *E.g.*, OECD, REPORT TO THE NEGOTIATING GROUP ON INTELLECTUAL PROPERTY, NEGOTIATING GROUP ON THE MULTILATERAL AGREEMENT ON INVESTMENT (MAI) 4 (1997), available at <http://www1.oecd.org/daf/mai/pdf/ng/ng9732e.pdf> (noting that some experts recommend excluding intellectual property rights entirely from the definition of “investment”).

348. *E.g.*, 35 U.S.C. § 261 (2012) (patents have attributes of personal property); *see also* Mercurio, *supra* note 26, at 878 (noting that it is “beyond doubt” that granted intellectual property rights are investments).

from the scope of the treaty.³⁴⁹ Alternatively, agreements could include language that states “[n]othing in this agreement shall affect the rights and obligations of any party to TRIPS or any other international intellectual property agreement; no party may bring an issue requiring adjudication of a TRIPS provision unless it has been previously determined to be in violation of TRIPS pursuant to the WTO.” This would be somewhat similar to existing exceptions in some agreements concerning either tax or environmental agreements.³⁵⁰ However, unlike these clauses, which are primarily conflict of law principles that state which agreement should prevail in the event of inconsistency, this proposal goes further to ensure that tribunals do not unnecessarily decide whether there is an inconsistency in the first instance. This is necessary to prevent commercial arbitrators from usurping the process for determining TRIPS compliance, which could lead to inconsistent judgments.³⁵¹

The above suggestions are strongly preferable to the draft text to address situations where there are competing agreements under CETA that is being used as a basis for the pending Transatlantic Trade and Investment Partnership (“TTIP”). In particular, the CETA draft states that if there is a potential for “overlapping compensation” or the other claim could have a “significant impact” on the arbitration claim, the tribunal shall “stay its proceedings,” or otherwise, the tribunal can continue the proceedings and simply take a separate proceeding “into account in its decision, order or award.”³⁵² However, this approach still not only gives a tribunal too much authority to impinge on another international agreement, but it also does not address the situation raised by Eli Lilly where no other proceeding has previously been initiated. This may often be the case with TRIPS claims because only governments can bring WTO disputes and governments seem circumspect in doing so. Moreover, WTO claims would not result in overlapping compensation both because investors have no standing to assert such claims, and because WTO proceedings are only intended to force

349. *E.g.*, NAFTA, *supra* note 36, arts. 2102, 2103.

350. *Id.* arts. 103, 2103; Korea-Chile Free Trade Agreement, Kor.-Chile, art. 20.3(3), Feb. 15, 2003, *available at* http://www.sice.oas.org/Trade/Chi-SKorea_e/ChiKoreaind_e.asp; *see also* Guillermo Aguilar Alvarez & William W. Park, *The New Face of Investment Arbitration: NAFTA Chapter 11*, 28 YALE J. INT’L L. 365, 390 (2003); Jennifer A. Heindl, *Toward a History of NAFTA’s Chapter Eleven*, 24 BERKELEY J. INT’L L. 672 (2006).

351. *E.g.*, Brooks E. Allen & Tommaso Soave, *Jurisdictional Overlap in WTO Dispute Settlement and Investment Arbitration*, 30 ARB. INT’L L. 7 (2014).

352. Draft CETA Investment Text, *supra* note 342, art. x-23; EUR. COMM’N TRADE, PUBLIC CONSULTATION ON MODALITIES FOR INVESTMENT PROTECTION AND ISDS IN TTIP, Question 7 (Mar. 2014) (citing CETA art. x-23), *available at* http://trade.ec.europa.eu/doclib/docs/2014/march/tradoc_152280.pdf.

countries to comply with WTO rules, but do not result in compensation.³⁵³ Accordingly, investment chapters should exclude from the scope of arbitrations any claims that challenge internationally agreed upon standards for state action.

2. *Limit the Scope of Investment Claims Based on International Agreements Such as TRIPS*

If intellectual property issues cannot be entirely excluded from investment arbitration disputes, the next best alternative is to cabin the most likely claims—expropriation and fair and equitable treatment claims. This section explains how to limit such claims and why existing proposals thus far are inadequate.

a) *Limit Expropriation Claims*

The optimal method of limiting challenges to domestic laws consistent with international intellectual property standards is to explicitly bar expropriation claims in this area. Technically, this is already recognized in existing agreements, including NAFTA. However, as the Eli Lilly case illustrates, that language is inadequate since parties may disagree on whether certain conduct is permissible under an international intellectual property agreement.

Canada has proposed that there is no indirect expropriation in the case of a decision by a court, administrative tribunal, or other governmental intellectual property authority limiting or creating an intellectual property right, except where “the decision amounts to a denial of justice or an abuse of right.”³⁵⁴ This would at first glance seem to easily bar claims like Eli Lilly’s without needing to evaluate whether there is a violation of a separate international agreement. However, a company, such as Eli Lilly, could claim a denial of justice or abuse of right. Although no prior tribunal has found similar facts to fit these circumstances, past expansive rulings suggest this is a possibility. Accordingly, any exception to expropriation for intellectual property rights should clarify that there is no denial of justice or abuse of right if there is a common law modification of laws that are retroactively applied. This would not only prevent the Eli Lilly situation, but would also make expropriation more in line with domestic taking law that does not

353. Understanding on Rules and Procedures Governing the Settlement of Disputes art. 3.7 Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, 33 I.L.M. 1226, 1241 (providing that compensation should only be awarded as a matter of last resort), available at http://www.wto.org/english/docs_e/legal_e/28-dsu.pdf.

354. Draft CETA Investment Text, *supra* note 342, art. X.11(5).

recognize a taking when courts simply apply slightly modified common law doctrine.³⁵⁵ Of course, there is no requirement that international expropriation must be consistent with domestic taking law. However, given that expropriation is a remedy only available to foreign investors, unless there is a sound policy reason to provide a broader scope of expropriation to only foreign investors, closer alignment in treatment of all investors seems most appropriate.

Another possibility is to bar expropriation claims based on intellectual property rights in a manner similar to expropriation claims based on taxation. For example, NAFTA states that tax measures may in some cases constitute expropriation, but imposes unique procedural requirements for asserting such a claim.³⁵⁶ In particular, before a claim can be adjudicated, both the country accused of expropriation as well as the investor's own country must decide whether there is an expropriation claim that is permitted to go forward.³⁵⁷ The idea of cabining expropriation claims based on domestic revocation of intellectual property rights is a sound one. To prevent potential inconsistent decisions, expropriation claims based on state action that is arguably inconsistent with TRIPS should be barred unless there is a finding of TRIPS inconsistency by a WTO panel. This would obviate inconsistent decisions and also allow TRIPS issues to be decided by arbitrators with expertise in WTO agreements, including TRIPS.

These proposals would be a significant improvement over the European Union's proposed language to clarify what types of regulatory action should not constitute indirect expropriation. Although the European Union shares a desire with many others to "avoid claims against legitimate public policy measures," its proposed clarification is no better than language in existing treaties.³⁵⁸ In particular it notes that in "rare circumstances," nondiscriminatory measures to protect "legitimate public welfare objectives such as health" can nonetheless constitute indirect expropriation if the impact of the measure "is so severe in light of its purpose that it appears manifestly excessive."³⁵⁹ This proposal introduces new language in need of interpretation, such as what would be "manifestly excessive" or "severe in light of its purpose." In addition, although it may seem fair to have a balance

355. *E.g.*, Vicki Been & Joel C. Beauvais, *The Global Fifth Amendment? NAFTA's Investment Protections and the Misguided Quest for an International "Regulatory Takings" Doctrine*, 78 N.Y.U. L. REV. 30, 80–82 (2003).

356. NAFTA, *supra* note 36, art. 2103(6).

357. *Id.*

358. EU Comm'n, Public Consultation on Modalities for Investment Protection and ISDS in TTIP, at 6–7 (Mar. 2014).

359. Draft CETA Investment Text, *supra* note 342, Annex X.11: Expropriation ¶ 3.

of interests, a tribunal of private arbitrators does the balancing, and in doing so the arbitrators are essentially second guessing the balance already determined by a nation.

b) Limit Fair and Equitable Treatment Claims

The best approach to cabining fair and equitable treatment claims would be to eliminate them altogether in cases where the agreement is solely between countries with strong legal systems. Although it may seem radical to jettison a traditional component of investment chapters, there are some existing chapters where tribunals have no authority to litigate such claims. These claims have posed the most significant intrusions into domestic regulatory authority and have resulted in inconsistent rulings. Moreover, this standard was initially intended to provide a remedy as a back up to the non-discrimination provision in the exceptional situation where the host country's political and legal systems disintegrate to the extent that investors cannot be adequately protected. There seems to be no need for this claim at all where domestic remedies exist. In addition, this would avoid the problem of unduly expansive rulings concerning fair and equitable treatment that the United States has tried, but failed to cabin in NAFTA.³⁶⁰

If fair and equitable treatment claims must remain within the scope of investment arbitrations, adding clear exceptions would be the next best alternative. For example, just as intellectual property rights denied or canceled under domestic law should never be considered expropriation, a similar clause could exist for fair and equitable treatment claims. In addition, as noted earlier with expropriation claims, it may be better to exclude any fair and equitable treatment claim based on state law denying or canceling an intellectual property right on substantive grounds unless that state law is found by a WTO panel to be inconsistent with TRIPS. Even if there were a TRIPS violation, there should not necessarily be a fair and equitable treatment claim. Many existing and pending agreements state that breach of a separate international agreement—including TRIPS—does not establish a

360. For example, after some broad interpretations of this standard under NAFTA, the NAFTA Free Trade Commission issued an interpretation that aimed to clarify that the standard be linked to customary international law to cabin rulings. NAFTA Free Trade Comm'n, Notes of Interpretation of Certain Chapter 11 Provisions, ¶ B (July 31, 2001). However, this was of little utility since tribunals simply interpreted customary international law broadly. E.g., Patrick Dumberry, *The Emergence of a Consistent Case Law: How NAFTA Tribunals Have Interpreted the Fair and Equitable Treatment Standard*, KLUWER ARB. BLOG (Oct. 30, 2013), <http://kluwarbitrationblog.com/blog/2013/10/30/the-emergence-of-a-consistent-case-law-how-nafta-tribunals-have-interpreted-the-fair-and-equitable-treatment-standard/>.

violation of the fair and equitable treatment standard.³⁶¹ Notably, this clause also does not state that compliance with another international agreement will immunize state action from being subject to such claims. Thus, additional language is necessary.

This could be accomplished by including language defining what qualifies as fair and equitable treatment. Although some agreements limit the term to the minimum standard pursuant to customary international law, such limits have clearly been inadequate in cabining intrusive claims. Accordingly the term could be stated to never exist simply because the legal or business environment has changed. This would importantly be helpful not only for the intellectual property issues that this Article focuses on, but also for all investor-state claims that have resulted in undue encroachment on domestic regulatory authority. In addition, an investment chapter could mandate that fair and equitable treatment claims must be based on whether a party made a specific representation to induce investment that created a legitimate expectation. Furthermore, it could mandate that there is never a legitimate expectation that laws will remain frozen in time. This would go farther than the current EU proposal that suggests that tribunals may consider whether a country made a specific representation relied upon by the investor.³⁶² In addition to requiring—rather than permitting—tribunals to consider specific representation, it may be important to define what constitutes such a representation. For example, Eli Lilly incorrectly believes that an intellectual property right granted by the state is a representation that the right can never be invalidated. Accordingly, it could be helpful to clarify that intellectual property rights issued by a nation are not representations of permanent validity.

IV. CONCLUSION

Eli Lilly's case against Canada exposes important policy problems with permitting investors to use investor-state arbitrations to challenge domestic intellectual property decisions. Although a tribunal should deny Eli Lilly's claims, investor-state tribunals often make broad and unpredictable rulings. Moreover, even if a panel rules properly, public health may still be compromised if other companies follow Eli Lilly's lead in challenging other domestic decisions concerning intellectual property rights. Although some are wisely beginning to question the wisdom of creating more opportunities

361. Draft, TPP, Investment chapter, *supra* note 179, art. II.6: Minimum Standard of Treatment.

362. See EU COMM'N TRADE, *supra* note 352, ques. 3.

through additional agreements, this Article hopes to provide a roadmap for how to combat likely claims in the thousands of existing agreements, as well as how to cabin claims in any future agreements.

IS COPYRIGHT INFRINGEMENT A STRICT LIABILITY TORT?

Patrick R. Goold[†]

ABSTRACT

Scholars and lawmakers routinely refer to copyright infringement as a strict liability tort. The strictness of copyright liability has long been criticized as immoral, inefficient, and inconsistent with usual tort doctrine. However, this Article questions whether copyright infringement really is a strict liability tort. It advances the thesis that copyright infringement in the United States is a fault-based tort, closely related to the tort of negligence. Using both doctrinal and economic methods, this Article explicates the role that fault plays in copyright infringement. Doing so not only demonstrates that copyright's liability rule is more normatively defensible than previously appreciated, but also provides a unique tort perspective on the nature of the fair use doctrine. By seriously engaging with the analytic question of whether liability for copyright infringement is strict or not, we highlight how the fair use analysis blends and confuses two separate issues: on one hand, did the defendant cause the plaintiff harm, and, on the other, was that harm justifiable? The Article concludes that, while no substantive changes need to be made to copyright's liability rule, judges ought to restructure the fair use analysis in order to keep these concepts distinct from one another.

© 2015 Patrick R. Goold

[†] IP Fellow at IIT Chicago-Kent College of Law. The author would like to thank dearly the following people for their advice and comments: Shyamkrishna Balganes, Christopher Buccafusco, Oren Bracha, Mark Gergen, Wendy Gordon, Justin Hughes, Edward Lee, Jake Linford, Kylie Pappalardo, Matthew Sag, Pamela Samuelson, Rebecca Tushnet, Richard Wright, Peter Yu. In addition, the author would like to thank the organizers and participants of the 2014 Santa Clara Works in Progress in Intellectual Property Workshop, the 2014 UC Berkeley Intellectual Property Scholars Conference, and the 2014 Michigan State Junior Scholars in Intellectual Property Conference. A special note of gratitude goes to the patient editors of the Berkeley Technology Law Journal. All mistakes are solely the responsibility of the author. Please send comments to pgould@kentlaw.iit.edu.

TABLE OF CONTENTS

I. INTRODUCTION	308
II. STRICT LIABILITY AND FAULT LIABILITY	312
A. THE DOCTRINE OF STRICT LIABILITY AND FAULT LIABILITY	312
1. <i>Strict Liability</i>	312
2. <i>Fault Liability</i>	314
a) State of Mind Fault	314
b) Standard of Conduct Fault	315
3. <i>Defenses</i>	319
B. THE ECONOMICS OF STRICT LIABILITY AND FAULT LIABILITY	321
1. <i>Economics Foundation</i>	322
2. <i>The Economic Goal of Tort Law</i>	323
3. <i>Strict Liability Rules</i>	324
4. <i>Negligence Rules</i>	325
5. <i>The Substantive Difference between Strict Liability and Negligence</i>	325
III. STRICT LIABILITY AND FAULT LIABILITY IN COPYRIGHT INFRINGEMENT	326
A. BASIC COPYRIGHT DOCTRINE	326
1. <i>The Prima Facie Case</i>	326
2. <i>Fair Use</i>	328
a) The Market Failure Approach	330
b) The Balancing of Public Interests Approach	331
B. EXISTING THEORIES OF COPYRIGHT INFRINGEMENT	333
1. <i>The Orthodox View</i>	333
2. <i>Professor Steven Hetcher's Fault Liability View</i>	335
a) Professor Hetcher's Argument	336
b) Critique	336
C. A DOCTRINAL REINTERPRETATION	338
1. <i>The Fault in Copyright Infringement</i>	338
a) A Blameworthy State of Mind?	339
b) A Failure to Comply with a Standard of Conduct	340
i) The Relationship of the Negligence Rule and the Fairness Rule	340
2. <i>The Harm in Copyright Infringement</i>	344
a) Isolating the Harm in Copyright Infringement	344
b) Harm and Fault in Fair Use	347
3. <i>Responding to the Orthodox View</i>	351
D. AN ECONOMIC REINTERPRETATION	353

1.	<i>The Economic Goal of Copyright Law</i>	353
2.	<i>A Strict Liability Rule in Copyright?</i>	354
3.	<i>A “Negligence Rule” in Copyright</i>	355
4.	<i>Incentives in Copyright: Strict Liability or Fairness?</i>	356
E.	CRITIQUES, COUNTER-ARGUMENTS AND A CAVEAT	357
1.	<i>Fair Copying Is Outside the Scope of the Right</i>	357
a)	Merit to the Critique.....	358
b)	Counter-Argument to the Critique.....	359
2.	<i>Fair Use is an Affirmative Defense</i>	361
a)	Fair Use as an Affirmative Defense.....	361
b)	The Procedural Role of Fault in Copyright Infringement	362
3.	<i>A Caveat: The Market Failure Approach to Fair Use</i>	364
IV.	RESTRUCTURING FAIR USE.....	365
A.	THE NORMATIVE DEFENSIBILITY OF COPYRIGHT INFRINGEMENT’S LIABILITY RULE	365
1.	<i>The Normative Critique</i>	366
a)	Inconsistency.....	366
b)	Inefficiency.....	367
c)	Immorality.....	368
2.	<i>Answering the Normative Critique</i>	368
a)	Inconsistency.....	368
b)	Inefficiency.....	369
c)	Immorality.....	370
3.	<i>Reforming Copyright Infringement as an Intentional Tort?</i>	370
a)	What Does Intent Mean in Copyright?	371
b)	Ought Copyright Infringement Be an Intentional Tort?	372
B.	THE FORMAL STRUCTURE OF COPYRIGHT INFRINGEMENT	374
1.	<i>Collapsing Harm and Fault</i>	374
2.	<i>The Burden of Proof</i>	377
a)	The Theory of Burden Shifting	377
b)	The Burden of Proving Harm and Fault	378
3.	<i>Solution: Restructuring the Fair Use Analysis</i>	379
V.	CONCLUSION.....	381

I. INTRODUCTION

Modern tort law has largely retreated from the principle of strict liability.¹ Although for many centuries, the common law imposed civil liability upon a defendant for harm that was not his fault, today the law typically requires that a defendant act intentionally, recklessly, or negligently before he will be held responsible for the consequences of his conduct.² For over a hundred years, jurists have largely applauded this transformation. The voices decrying strict liability come from the greatest figures of common law jurisprudence, such as Oliver Wendell Holmes—who argued that strict liability would wastefully deter productive activity³—to the foremost minds of contemporary legal thought, who argue that holding someone responsible without fault is potentially immoral⁴ and potentially inefficient.⁵ This evolution resulted in the situation where strict liability exists “at the margins of tort”⁶ applicable only in “a few special situations,”⁷ and a belief that it is a “mediaeval”⁸ concept that simply “does not fit” within the greater body of private law.⁹

1. See generally MORTON J. HORWITZ, *THE TRANSFORMATION OF AMERICAN LAW 1780–1860*, at 97–101 (1977); see also G. EDWARD WHITE, *TORT LAW IN AMERICA* 244–90 (1980); G. EDWARD WHITE, *The Unexpected Persistence of Negligence, 1980–2000*, 54 *VAND. L. REV.* 1337, 1344–46 (2011). But see Robert L. Rabin, *The Historical Development of the Fault Principle: A Reinterpretation*, 15 *GA L. REV.* 925 (1981).

2. See, e.g., DAN B. DOBBS, *THE LAW OF TORTS* § 342 (2008) (after 1841, “negligence or intentional invasions would thereafter become the normal basis for tort liability”); Cornelius J. Peck, *Negligence and Liability Without Fault in Tort Law*, 46 *WASH. L. REV.* 225, 225 (1971) (“It is frequently assumed that with a few exceptions the principles of negligence comprise the field of tort law, and that fault is the most common basis for determining liability for harmful conduct.”).

3. OLIVER WENDELL HOLMES, JR., *THE COMMON LAW* 95 (1881) (“As action cannot be avoided, and tends to the public good, there is obviously no policy in throwing the hazard on what is at once desirable and inevitable upon the actor.”); see also James Barr Ames, *Law and Morals*, 22 *HARV. L. REV.* 97, 99 (1908) (“The ethical standard of reasonable conduct has replaced the unmoral standard of acting at one’s peril.”).

4. See, e.g., Jules Coleman, *Moral Theories of Torts: Their Scope and Limits, Part I*, 1 *LAW & PHIL.* 371, 374 (1982) (“[T]he substitution of fault for causation marked an abandonment of the immoral standard of strict liability under Trespass (which, after all, imposed liability without regard to fault) in favor of a moral foundation for tort law based on the fault principle.”); ERNEST WEINRIB, *THE IDEA OF PRIVATE LAW* 170–203 (1995).

5. See, e.g., ROBERT COOTER & THOMAS ULEN, *LAW AND ECONOMICS* 208–11 (6th ed. 2012) (describing how, without a defense of contributory negligence, strict liability gives the victim inefficient incentives to take care).

6. JOHN C.P. GOLDBERG & BENJAMIN ZIPURSKY, *THE OXFORD INTRODUCTIONS TO U.S. LAW: TORTS* 265 (2010).

7. *Id.* at 266.

8. *Read v. J. Lyons & Co.*, [1945] K.B. 216 at 229 (Eng.) (Scott L.J.).

9. GOLDBERG & ZIPURSKY, *supra* note 6, at 267.

For several decades, scholars have tried to provide a plausible normative justification for holding individuals liable even when their actions are without fault.¹⁰

As strict liability is typically seen as the exception, not the rule, intellectual property scholars have become increasingly concerned about the state of copyright law. Copyright infringement, according to most judges and commentators, is a strict liability tort.¹¹ A plaintiff can establish a prima facie case of direct infringement merely by showing that a defendant copied his protected work and that this resulted in the production of a substantially similar work.¹² As there is no requirement on the plaintiff to show how the defendant behaved intentionally, recklessly, or even negligently, it is commonly said that “innocence is no defense to an action for copy-right infringement.”¹³ This situation has struck many as normatively untenable. Over seventy years ago, Judge Learned Hand worried that the application of strict liability in copyright was “harsh” and worthy of hesitation.¹⁴ More recently, academicians have maintained that exposing copyright defendants to strict liability is immoral, inefficient, and inconsistent with the standard tort practice of only holding liable those defendants who have acted wrongfully. To remedy this situation, a number of scholars have proposed that copyright reject strict liability in favor of a fault liability rule. In their vision, copyright law would be improved if it only imposed liability on those defendants who copy intentionally, recklessly, or negligently.¹⁵

10. See JAMES GORDLEY, *FOUNDATIONS OF PRIVATE LAW* iii (2006) (explaining the Late Scholastic scholars’ failed attempt to provide a rationale for strict liability inherited from Roman law); Richard A. Epstein, *A Theory of Strict Liability*, 2. J. LEGAL STUD. 151 (1973); Tony Honoré, *Responsibility and Luck: the Moral Basis of Strict Liability*, 104 L.Q. REV. 530 (1988).

11. See *infra* notes 128 & 129.

12. See *infra* pp. 327–28 and accompanying footnotes.

13. 2 PAUL GOLDSTEIN, *GOLDSTEIN ON COPYRIGHT* § 8.1 n.1 (3d ed. 2014).

14. *Barry v. Hughes*, 103 F.2d 427, 427 (2d Cir. 1939) (per curiam) (“It has been held that one who copies from a plagiarist is himself necessarily a plagiarist, however innocent he may be, but that would be a harsh result, and contrary to the general doctrine of torts. . . . We should hesitate a long while before holding that the use of material, apparently in the public demesne, subjected the user to damages, unless something put him actually on notice.” (internal citation omitted)); see also *De Acosta v. Brown*, 146 F.2d 408, 413 (2d Cir. 1944) (Hand, J., dissenting) (“Ordinarily an act does not become a wrong, when to make it so, one must resort to consequences arising from it in the actual sequence of events which reasonable persons would not anticipate. . . . I can see no reason why the ordinary rule of liability for torts should not apply to copying a copy . . .”); see also *Shapiro, Bernstein & Co. v. H.L. Green Co.*, 316 F.2d 304, 308 (2d Cir. 1963) (noting the “harshness of the principle of strict liability in copyright law”).

15. See, e.g., Kent Sinclair, Jr., *Liability for Copyright Infringement—Handling Innocence in a Strict-Liability Context*, 58 CALIF. L. REV. 940 (1970); Dane S. Ciolino & Erin A. Donelon,

However, despite the widespread and orthodox belief that copyright infringement is a strict liability tort, this characterization is questionable. A number of articles quickly classify liability for copyright infringement as strict, and then proceed with haste to the normative question of whether that state of affairs is desirable.¹⁶ Sadly, there is precious little discussion that seriously engages with the positive question of whether copyright infringement in the United States actually is based on strict liability.¹⁷ In an attempt to rectify the lack of descriptive theory in copyright law, this Article tries to answer the question in an analytically rigorous fashion. In doing so, the Article demonstrates that this issue is much more complicated than previous scholars have appreciated. Moreover, contrary to the dominant view of copyright infringement, this Article advances the thesis that copyright infringement is in fact a fault-based tort. In a nutshell, copyright infringement is not a strict liability tort because it does not hold the defendant liable simply on the basis that he infringed a right of the plaintiff. In addition, it must be shown that the defendant's copying was wrongful. The fair use doctrine exists, in part, to exculpate defendants who infringe a plaintiff's copyright but who do so in socially beneficial ways. Only those who infringe copyright unfairly, and who therefore wrongfully impose negative consequences upon the rest of society, are held liable.

At which point, one might ask: why does this matter? Even if one assumes the thesis presented here is correct, are not strict liability or fault

Questioning Strict Liability in Copyright, 54 RUTGERS L. REV. 351, 419–20 (2002) (arguing that lack of intent ought to be a defense to copyright infringement); Steven Hetcher, *The Kids Are Alright: Applying a Fault-Liability Standard to Amateur Digital Remix*, 62 FLA. L. REV. 1275 (2010) (arguing for a fault liability regime in relation to user-generated content); Assaf Jacob & Avihay Dorfman, *Copyright as Tort*, 12 THEORETICAL INQUIRIES IN LAW 59 (2011) (ARGUING that copyright infringement ought to adopt intentional, negligence, and strict liability rules in different contexts); Jacqueline D. Lipton, *Cyberspace, Exceptionalism, and Innocent Copyright Infringement*, 13 VAND. J. ENT. & TECH. L. REV. 767 (2011) (suggesting several ways in which innocence could result in findings of no liability); Tony Evans, *“Safe Harbor” for the Innocent Infringer in the Digital Age*, 50 WILLAMETTE L. REV. 1 (2013) (describing a DMCA-like safe harbor for direct unintentional infringement).

16. See, e.g., Ciolino & Donelon, *supra* note 15, at 356 (copyright infringement requires no “scienter, intent, knowledge negligence, or similar culpable mental state. On the contrary, liability for civil copyright infringement is strict”); Lipton, *supra* note 15, at 768 (“Historically, copyright infringement claims have been litigated on a strict liability basis.”); Evans, *supra* note 15, at 4 (referring to the “strict liability nature of copyright infringement that applies generally in all cases”). As will be shown, the issue is far more complex than such statements suggest. These statements overlook the fact that fault in law does not refer only to a defendant's subjective mental state. See *infra* pp. 314–19.

17. One exception comes from the work of Professor Steven Hetcher. See *infra* pp. 336–38 and accompanying footnotes. Similar lines of inquiry are also emerging in patent law. See Saurabh Vishnubhakat, *An Intentional Tort Theory of Patents*, FLA. L. REV. (forthcoming).

liability simply names? Characterizing liability as strict or not strict does not actually affect the underlying doctrine. Nevertheless, the problem arises in this context when we consider the normative debate that surrounds copyright's liability rule. As previous copyright scholars have paid little attention to the complex analytic question, they have erroneously characterized copyright as a strict liability tort and then proceeded to demonstrate why strict liability is normatively unattractive in this context. By arguing that copyright infringement is not strict, this Article demonstrates that much of the handwringing is misplaced. As copyright is already based upon fault, it is less inconsistent, inefficient, and immoral than previously supposed.

Furthermore, because previous authors have skipped over the complex analytic question and rushed to the normative one, they have missed an even more pressing concern: the formal structure of copyright infringement is a mess! Asking the question "is liability in copyright strict or not?" provides a unique tort perspective on the nature of copyright infringement generally, and the fair use doctrine in particular. The key insight this inquiry reveals is that the fair use doctrine currently blends and confuses two separate inquiries, namely: (a) has the defendant caused the plaintiff harm and (b) was that harm justifiable? Sadly, this conflation is largely pernicious. As will be elaborated upon, it not only causes judges to fit cases of "no fault" into the language of "no harm," thus prejudicing defendants with legitimate "no fault" claims, but it also results in poorly assigned burdens of proof. Therefore, after showing that copyright is a fault-based tort where the standard of fault is normatively defensible, the Article demonstrates how judges could restructure the fair use analysis so that these concepts are separated from one another.

Part II of this Article uses both doctrinal and economic methods to demonstrate the distinction between strict liability and fault liability rules. Part III applies this framework to copyright infringement. Doing so demonstrates two things: firstly, the question of whether copyright is strict or fault-based is far more complex than previously appreciated, and secondly, there is an arguable case that copyright infringement is a fault-based tort. Once these analytic points are developed, Part IV enters into the normative debate surrounding copyright infringement's liability rule. If copyright infringement is already based on fault, then the system is more tenable than some have previously appreciated. Nonetheless, while courts need not alter the substance of the liability rule in place, they must pay more attention to the formal structure of this rule. In particular, they ought to distinguish more carefully the two separate concepts of harm and fault that are embedded in the fair use analysis. This Part offers a way in which such separation could feasibly be accomplished. Part V concludes.

II. STRICT LIABILITY AND FAULT LIABILITY

This Part shall compare strict liability and fault liability rules. The first Section is doctrinal. It explains the legal difference between these two types of liability. The second Section is economic and functional. It explains the utilitarian goal the law attempts to serve and illustrates how both strict liability and fault liability rules achieve that goal.

A. THE DOCTRINE OF STRICT LIABILITY AND FAULT LIABILITY

Before a court will hold a defendant responsible, the plaintiff must demonstrate that he has a legitimate *prima facie* case. To do so, he must prove the existence of several factual conditions. These conditions vary depending on the type of liability rule the law adopts. Generally speaking, tort uses two forms of liability rule: strict liability and fault liability. This Section demonstrates the conditions that must be established before a defendant will be held liable under a strict liability and under a fault liability rule.

1. *Strict Liability*

Strict liability is liability imposed when a defendant infringes the legal right of another person.¹⁸ As legal rights differ in character, we find there are two different categories of strict liability: conduct-based strict liability and harm-based strict liability.

Autonomy rights confer upon the right holder a broad power to control an object.¹⁹ This right of control is infringed whenever the defendant engages

18. This definition might seem a little unorthodox. More commonly strict liability is defined as liability imposed upon an individual whose conduct causes harm to the plaintiff. *See, e.g.*, GOLDBERG & ZIPURSKY, *supra* note 6, at 90 (“Under a regime of strict liability, an actor who causes harm to another is held liable simply by virtue of causing harm.”); JULES COLEMAN, *RISKS AND WRONGS*, 212–34 (2d ed. 2002). Thus, strict liability is considered liability based on causation, while fault liability is liability based on causation plus fault. However, such a definition would apparently not cover cases such as trespass to land, where liability is imposed regardless of whether the defendant’s conduct “caused” some form of “harm.” In trespass, conduct alone seems to be the touchstone for liability, not causation of harm. Perhaps a better definition therefore is simply that strict liability is liability imposed “regardless of fault.” *See* PETER CANE, *RESPONSIBILITY IN LAW AND MORALITY* 82 (2002). However, such a negative definition does not actually tell us what is the justification for liability in such cases; it merely tells us that fault is not the relevant justification. Therefore, this article prefers to define strict liability as rights-infringement. This view is supported by recent analytic theory of strict liability by Professor Greg Keating. *See* Gregory Keating, *Strict Liability Wrongs*, in *THE PHILOSOPHICAL FOUNDATIONS OF THE LAW OF TORTS* 295 (John Oberdiek ed., 2014) [hereinafter Keating, *Wrongs*]; Gregory Keating, *Nuisance as a Strict Liability Wrong*, 4 J. TORT LAW 1 (2012) [hereinafter Keating, *Nuisance*].

19. Keating, *Wrongs*, *supra* note 18, at 296–300.

in conduct that is antithetical to that control. The right holder need not suffer any real world harm before the right is invaded. The only “harm” he need suffer is a legal one (i.e., the lost power to control the object).²⁰ For example, property rights are typically autonomy rights. One’s property right in land is infringed if someone else enters the land without permission. There need not be any real world harm flowing from the entry before the right is infringed.

Autonomy rights are protected by conduct-based strict liability rules.²¹ These rules attach liability to a defendant who engages in a form of proscribed conduct.²² Importantly, the plaintiff need not demonstrate how this volitional conduct caused a harmful outcome before liability is imposed. The classic example of this is the tort of trespass to land. As the right to exclude is an autonomy right, the law imposes liability upon the defendant who volitionally enters the land even when that entry is not harmful.²³

Alternatively, some rights do not confer broad powers of control, but instead only the right to maintain an object in a certain condition.²⁴ These rights can be infringed only if the defendant’s actions cause some real world harmful consequence; harm, not the lost power to control, grounds liability in this instance. For example, one has a right to physical health and being “whole in body and mind.”²⁵ As the right is to maintain one’s health, the right can only be infringed if the defendant’s action causes the victim’s health to deteriorate or worsen.

Such rights are protected by harm-based strict liability rules.²⁶ These rules attach liability to a defendant who volitionally engages in a form of proscribed conduct that causes a harmful outcome. For example, products

20. This may also be referred to as a “normative” loss. *See, e.g.*, John C.P. Goldberg & Benjamin Zipursky, *Torts as Wrongs*, 88 TEX. L. REV. 917, 956 (2010).

21. Keating, *supra* note 19.

22. PETER CANE, *THE ANATOMY OF TORT LAW* 45 (1997); *see also* Shyamkrishna Balganesh, *The Obligatory Structure of Copyright Law: Unbundling the Wrong of Copying*, 125 HARV. L. REV. 1664, 1682 (2012).

23. Trespass to land is sometimes mistakenly called an intentional tort. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 158 (calling trespass an “intentional intrusion on land”). This mistake comes from confusing the concepts of volition, deliberateness, and intentionality. *See* CANE, *supra* note 22, at 32–33; *see also* DOBBS, *supra* note 2, § 51 (“Since the intent required to show a trespass is only an intent to enter land, and since that intent might be wholly innocent, the rules may sometimes impose a limited kind of strict liability.”).

24. Keating, *Wrongs*, *supra* note 18, at 296–300.

25. *See, e.g.*, CANE, *supra* note 22, at 67 (discussing one’s physical interest in good health).

26. CANE, *supra* note 22, at 47–49 (discussing “outcome-based” strict liability). *See also* Balganesh, *supra* note 22.

liability adopts a harm-based strict liability rule. Because the consumer has a right to bodily health, liability is imposed on the defendant who manufactures a defective product (the proscribed conduct) that in turn causes the consumer some physical injury (the harmful outcome).²⁷

Importantly, however, neither conduct-based strict liability nor harm-based strict liability is conditioned upon fault. The defendant's infringement of the plaintiff's right need not be wrongful for liability to be imposed. Indeed, strict liability imposes liability even when the defendant's conduct is deemed rightful and a good thing for society. For example, abnormally dangerous activities are governed by strict liability rules.²⁸ Society acknowledges that engaging in abnormally dangerous activity is often a good thing. Sometimes we must engage in abnormally dangerous activities such as crop dusting or dynamite blasting for the overall benefit of society.²⁹ Accordingly, engaging in such conduct is not considered wrongful. Nevertheless, despite the fact that it is not wrong to engage in this conduct, the law still makes the person who does so liable to the plaintiff if it results in an infringement of a right.

2. *Fault Liability*

Fault liability rules are harm-based strict liability rules with one additional element: fault.³⁰ Liability is not imposed solely upon rights-infringement. A defendant is only held responsible if he has engaged in the proscribed conduct that in turn causes a harmful outcome, and when that conduct is deemed to be wrongful.³¹ Tort law recognizes two categories of wrongful conduct.³² Firstly, a defendant's conduct is wrongful if he acts with a blameworthy state of mind. Secondly, a defendant's conduct is wrongful if it fails to live up to a standard that the law expects. The following subsections explain these two different categories of fault.

a) State of Mind Fault

Fault may be established by demonstrating that the defendant acted with a blameworthy state of mind.³³ This is most commonly achieved by proving

27. RESTATEMENT (SECOND) OF TORTS § 402; *see generally* DOBBS, *supra* note 2, § 354.

28. *Id.* § 519(a).

29. *Id.* § 520.

30. *See* COLEMAN, *supra* note 18, at 212; *see also* GOLDBERG & ZIPURSKY, *supra* note 6, at 90–91.

31. COLEMAN, *supra* note 18, at 212.

32. CANE, *supra* note 18, at 78 (2002) (“Legal fault consists either of a failure to comply with a specified standard of conduct, or of failure to comply with a specified standard of conduct accompanied by a specific state of mind.”)

33. COLEMAN, *supra* note 18, at 217–18.

that the defendant acted intentionally. By acting intentionally, the defendant engaged in the conduct with the aim of causing the harmful outcome.³⁴ Note, this is not the same as acting volitionally. Conduct is volitional when engaged in voluntarily;³⁵ conduct is intentional when that conduct is engaged in to cause some harmful consequence. For example, battery is an intentional tort.³⁶ To prove battery the plaintiff must show that the defendant volitionally touched the defendant, that this touching was harmful to the plaintiff (either by showing physical injury or that the touching was “offensive”), and that the defendant intended that such contact be harmful.³⁷ Hence, a leading treatise on tort states that, in an action for battery, “[a]n intent to cause actual harm is sufficient intent but not a necessary one. It is enough that the defendant intends bodily contact that is offensive.”³⁸

In other cases, the plaintiff can demonstrate the defendant’s blameworthy state of mind by demonstrating that the defendant caused the harm recklessly (i.e., that he consciously disregarded an unreasonable risk), fraudulently (i.e., that he intended to deceive the plaintiff), or maliciously (i.e., that he acted with bad motives).³⁹ But these states of mind are less commonly required as the basis for liability.

b) Standard of Conduct Fault

Alternatively, a plaintiff may prove fault by demonstrating that the defendant’s conduct simply failed to live up to a standard that the law expects of him. This is most commonly achieved by establishing that the defendant caused the harmful outcome negligently.⁴⁰ In such cases, the standard expected is that individuals will conduct themselves reasonably. A defendant’s conduct is negligent if he failed to act in the manner of a

34. CANE, *supra* note 22, at 32–33.

35. *Id.* at 29–32.

36. RESTATEMENT (SECOND) OF TORTS § 13.

37. *Id.*

38. DOBBS, *supra* note 2, § 8.

39. *See generally* CANE, *supra* note 22, at 33–36.

40. *Id.* at 36 (“[D]eliberate, intentional and reckless conduct alike may attract tort liability for negligence if the conduct satisfies the definition of negligence, which is in terms of failure to attain a certain standard. In this way, the concept of negligence in tort law is rather different from the non-legal concept of carelessness, which implies inadvertence or lack of deliberation.”); RESTATEMENT (SECOND) OF TORTS § 282 (stating that the fundamental question in negligence law is whether conduct falls below the “standard established by the law for the protection of others against unreasonable risk of harm”); COLEMAN, *supra* note 18, at 217 (“An action is at fault when it fails to measure up to the relevant standard of conduct.”); COLEMAN, *supra* note 18, at 332 (“If she has failed to take reasonable care, then her conduct falls below an *objective* standard of conduct.”).

“reasonable person.”⁴¹ Judging a defendant’s conduct by a reasonableness standard is often referred to as the “negligence rule.”

When discussing negligence, four points must be clear. First, unlike intentional, reckless, fraudulent, or malicious conduct, negligence does not depend upon the defendant’s mental state. All that matters is the factual relationship between the defendant’s conduct and the legal standard. Hence, a defendant who unintentionally engages in unreasonable conduct is just as negligent as a defendant who intentionally engages in unreasonable conduct. This principle informed Professor Henry Terry’s statement that negligence is “conduct, not a state of mind.”⁴²

Second, the negligence rule is distinguishable from the “tort of negligence.”⁴³ The tort of negligence is a cause of action that sanctions a defendant for taking unreasonable risks that cause harmful accidents.⁴⁴ The negligence rule, by contrast, is not a cause of action, but the standard by which the conduct is judged.⁴⁵ The negligence rule is therefore applied in the tort of negligence, but equally the negligence rule is also applied in other causes of action, such as private nuisance⁴⁶ or defamation.⁴⁷

Third, the concept of “reasonableness” has no precise definition. It is a flexible standard that changes depending upon the facts of the case. Nevertheless, reasonableness is most commonly explained in consequentialist terms.⁴⁸ Whether conduct is reasonable depends upon whether it creates

41. DOBBS, *supra* note 2, § 117; *Vaughan v. Menlove*, (1837) 132 Eng. Rep. 490 (C.P.); 3 Bing. N.C. 468 (holding a defendant liable although he could not have done any differently due to a disability).

42. Henry Terry, *Negligence*, 29 HARV. L. REV. 40, 40 (1915); *see also* CANE, *supra* note 18, at 111 (2002) (law “recognises that failure to comply with standards of conduct can be culpable regardless of choice”).

43. CANE, *supra* note 22, at 36.

44. *Id.*

45. *Id.*

46. RESTATEMENT (SECOND) OF TORTS § 822 (requiring “unreasonable” interference with land).

47. *Id.* § 558 (requiring “fault amounting at least to negligence”).

48. I certainly do not mean that a deontological interpretation of reasonableness is impossible. But even deontological scholars have noted that negligence is usually discussed in consequentialist terms. *See e.g.* Heidi M. Hurd, *The Deontology of Negligence*, 76 B.U. L. REV. 249 (1996) (“It should be a great puzzle to those who consider themselves deontologists that the concept of negligence is most often, and certainly most clearly, defined in the moral language common to consequentialists.”); *see also* George Fletcher, *Fairness and Utility in Tort Theory*, 85 HARV. L. REV. 537 (1972) (associating the “reasonableness paradigm” with the dominant instrumentalist and utilitarian philosophy in U.S. tort scholarship).

greater benefit or cost for society.⁴⁹ In these terms, the reason why negligent conduct is deemed wrongful is that it forces negative consequences upon the rest of society. In the tort of negligence, this takes the form of increased risk of harmful accidents.

Fourth, the types of cost and benefits that are salient to the consequentialist balancing determination are often different across different torts. For example, in the tort of negligence, where the proscribed conduct is risk-taking, the relevant cost is the increased probability of an accident and the benefit is the reduction of resources spent on avoiding the accident. As Judge Learned Hand explained, in the tort of negligence, it is reasonable to take risks where the cost of precaution exceeds the expected accident costs.⁵⁰ Alternatively, it is unreasonable to take risks where the expected accident costs exceed the cost of precaution. But this formula clearly does not apply to the negligence rule as it appears in other causes of action. For example, in certain circumstances, the tort of private nuisance may also be said to adopt a negligence rule.⁵¹ A private nuisance is a substantial and unreasonable interference with the use and enjoyment of land. Whether a defendant's conduct is reasonable depends, in some interpretations, upon a consequentialist balancing test. But here the relevant costs and benefits of the action that are weighed are not the cost of precaution and benefit of avoiding an accident. Instead they are the gravity of the interference and the social utility of the activity.⁵² Because the causes of action govern different types of conduct, the utilitarian balancing calculus necessarily is based on different parameters.

In addition to these points, we must also distinguish cases where the law uses a standard to judge the wrongfulness of a defendant's conduct from

49. John C.P. Goldberg and Benjamin Zipursky, *Torts as Wrongs*, *supra* note 20, at 936 (“For utilitarians, it is said, the wrongfulness of conduct hinges on the probability that the conduct will produce net disutility (more pain than pleasure).”).

50. *United States v. Carroll Towing Co.*, 159 F.2d 169, 173 (2d Cir. 1947) (holding that liability is imposed when the cost of the burden is less than the gravity of injury multiplied by probability of it occurring).

51. Cane, *The Anatomy of Tort Law*, *supra* note 22, 145 (“The requirement of unreasonableness is practically equivalent to a requirement of negligence: the interference with use and enjoyment must have been foreseeable, and it must be greater than it is reasonable to expect P to put up with.”); William M. Landes & Richard A. Posner, *The Economic Structure of Tort Law* 41–53, 45 (1987) (“In most nuisance cases the standard is not strict liability but reasonableness, equivalent to nonnegligence.”) In the economic interpretation of tort law, nuisance adopts a simple negligence rule but it is assumed that harm is certain, rather than merely probable. As a result, judges do not compare the costs of precaution against the ex ante expected cost of harm, but simply against the ex post total cost of harm.

52. RESTATEMENT (SECOND) OF TORTS § 826(a).

cases where the law uses a standard simply to define the scope of a legal right. This is a subtle distinction, and one that is under-theorized, but is best demonstrated by private nuisance. In some interpretations, land owners are deemed to have a right only to the “reasonable use and enjoyment of land.”⁵³ The land owner has a right to maintain an object in a certain condition (i.e., that use and enjoyment be maintained at a reasonable level). Thus, if the court asks, “did the defendant’s actions cause the plaintiff’s use and enjoyment to drop below a reasonable level?” then the court is simply asking whether the defendant infringed a right.⁵⁴ By contrast, if, after deciding that the defendant’s right to reasonable use and enjoyment has been infringed, the court proceeds to ask “did the defendant behave reasonably?” (where reasonable conduct is defined as producing greater benefits than cost), then the court is asking whether the defendant’s conduct was wrongful.⁵⁵ Thus, in both cases, the law employs a standard to aid its determination, but in the former case, the court uses that standard to determine whether a right was infringed, and in the latter case, it uses a standard to determine whether the rights-infringement was wrongful.⁵⁶

Finally, some more general points on the distinction between state of mind fault and standard of conduct must be highlighted. State of mind fault is often referred to as fault in the actor.⁵⁷ By contrast, standard of conduct fault is often referred to as fault in the action.⁵⁸ This refers to the fact that the fault in the former case is internal to the defendant, whereas in the latter case the fault is in the defendant’s external actions. Similarly, state of mind fault is subjective (i.e., its existence depends on what the defendant was thinking at the time). On the other hand, standard of conduct fault is objective.⁵⁹ The existence of this fault does not depend on the actor’s personal point of view; all that matters is the factual relationship between his conduct and the standard.⁶⁰

53. Keating, *Nuisance*, *supra* note 18, at 14.

54. *Id.* at 35–42; Richard Wright, *Private Nuisance Law: A Window on Substantive Justice*, in ANDREW ROBERTSON & DONAL NOLAN, *RIGHTS AND PRIVATE LAW* 491, 502–08 (2011).

55. *Id.* at 35–42; WILLIAM M. LANDES & RICHARD A. POSNER, *ECONOMIC STRUCTURE OF TORT LAW* 39 (1987) (calling the reasonable use nuisance rule and a negligence rule “essentially equivalent.”).

56. Keating, *Nuisance*, *supra* note 18, at 39 (finding such balancing “akin to a judgment of fault in negligence”); *id.* at 40 (finding the balancing test in nuisance to be similar to the “application of the Hand formula with probability dropped out, because the harm is certain to occur”).

57. COLEMAN, *supra* note 18, at 217.

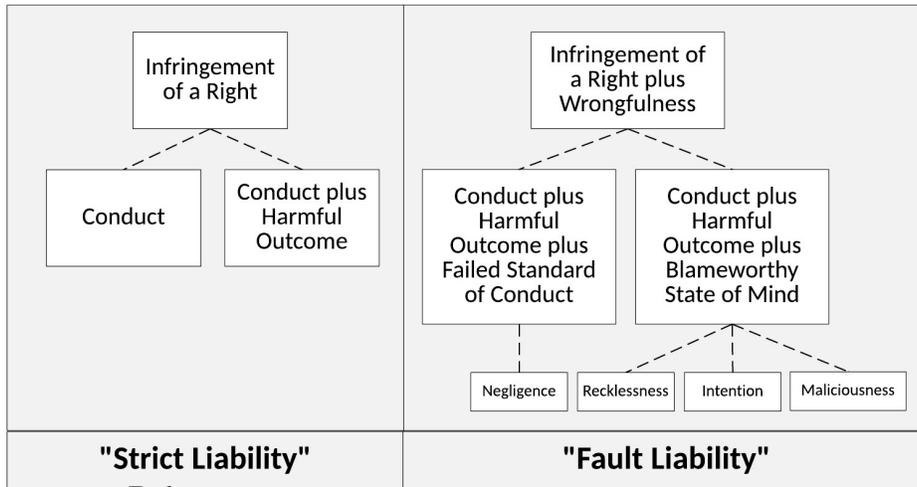
58. *Id.*

59. *See id.* at 225, 228.

60. *See* Terry, *supra* note 42, at 40.

Diagram 1 summarizes the difference between the elements of the prima facie case under strict liability rules and fault liability rules. The left side of the diagram represents strict liability rules, demonstrating that liability is imposed only upon the infringement of a right, which sometimes require proof only of conduct and sometimes requires additional proof of harmful outcome. The right side of the diagram represents fault liability rules and demonstrates that such liability is conditioned not only upon rights-infringement but also on wrongfulness, where wrongfulness is understood as either the failure to comply with a standard of conduct or acting with a blameworthy state of mind.

Diagram 1



3. Defenses

Once the plaintiff has established the elements of the prima facie case, the defendant is considered responsible for the accident as an initial matter. He then has the opportunity to exculpate himself by introducing affirmative defenses. The distinction between strict liability and fault liability can also be demonstrated by examining the defenses available under each liability rule.

We must first begin by separating three classes of affirmative defenses: plaintiff fault, justification, and excuse. Plaintiff fault defenses assert that the defendant should not be held liable because the plaintiff was at fault for his injury.⁶¹ The most common example of this is the contributory negligence

61. COLEMAN, *supra* note 18, at 227 (“an injurer is strictly liable but is given the opportunity to defeat his liability by showing the plaintiff himself is at fault”); Richard A. Epstein, *Defenses and Subsequent Pleas in a System of Strict Liability*, 3 J. LEGAL STUD. 165 (1974) (in addition to defenses of assumption of risk and plaintiff trespass, Epstein also believes the absence of causation is also a valid defense in strict liability actions; this Article takes the

defense.⁶² If the defendant can prove that the plaintiff's negligence contributed towards his injury, then the defendant will be exculpated. Other examples include where the plaintiff has voluntarily assumed the risk,⁶³ or, in products liability cases, where the plaintiff has altered or misused the product, resulting in his injury.⁶⁴

Justifications assert that, although the defendant has caused the plaintiff some harm, this conduct was not wrongful.⁶⁵ Instead, causing harm in this scenario was the right thing to do, and perhaps something the law aims to encourage. Classic examples include self-defense in battery cases⁶⁶ or truth in defamation cases.⁶⁷ Even though reasonable acts of self-defense may cause physical harm, and unfavorable published statements may cause reputational harm, the law takes the view that an individual can rightly engage in this conduct in certain situations.⁶⁸

Unlike justifications, excuses do not assert that the defendant's conduct was rightful. Instead, excuses are assertions that the defendant's conduct was understandable given his personal condition and, therefore, he is not personally blameworthy.⁶⁹ For example, in certain circumstances, the defenses of mental disability, infancy, and mistake exist to exculpate the defendant from tort liability. Unlike justifications, which focus on whether the defendant's actions were objectively wrongful or not, excuses focus on the subjective characteristics of the defendant.⁷⁰ Excuses are less commonly

view that absence of causation is not an affirmative defense but a claim that the prima facie case has not yet been established).

62. See generally DOBBS, *supra* note 2, § 199.

63. See generally DOBBS, *supra* note 2, § 211.

64. See generally DOBBS, *supra* note 2, § 370.

65. DOBBS, *supra* note 2, § 69 ("When a judge believes the defendant's harmful act was justified, the judge believes that people in general can rightly act as the defendant did."); JAMES GOUDKAMP, TORT LAW DEFENCES 76 (2013) ("Justificatory defences have been defined as defences that enable the defendant to escape from liability because, in committing a tort, the defendant acted reasonably.").

66. RESTATEMENT (SECOND) OF TORTS § 65.

67. *Id.* § 581A.

68. Admittedly, whether truth is an affirmative defense or, alternatively, whether falsity is an element of the prima facie claim is a doctrinally uncertain issue. See DOBBS, *supra* note 2, § 410 (noting that although the "mainstream common law thus recognizes truth as an affirmative defense . . . [.] Constitutional decisions have shifted the burden of proof on the issue of truth or falsity in cases involving certain public officials, public figures, or public-concerning issues . . .").

69. DOBBS, *supra* note 2, § 69 (excuses "assert that the defendant's conduct was understandable given his personal condition and that he is not personally blameworthy for matters not within his control. Excuses focus on subjective mental or psychological characteristics of the actor."); GOUDKAMP, *supra* note 65, at 83–85.

70. DOBBS, *supra* note 2, § 69.

available in tort law than justifications. This reflects the fact that the bulk of tort law deals with objective, not subjective, standards of liability.⁷¹

Crucially, the only class of defense available under a strict liability rule is plaintiff fault. Assumed risk, contributory negligence, and, in the case of products liability, unforeseen misuse and modification are the common methods of exculpation. However, as justification and excuse are not admissible affirmative defenses, strict liability is said to be liability “not defeasible by either excuse or justification.”⁷² By contrast, justification and excuse are admissible affirmative defenses under fault liability rules.

The reason for this distinction between strict liability and fault liability defenses is clear. Justifications and excuses both assert that the defendant was not at fault.⁷³ Justifications assert that the defendant’s conduct was objectively not wrongful, and therefore there is no fault in the action.⁷⁴ Excuses assert that, although the defendant’s conduct was wrongful, the individual is not morally blameworthy for the action; there is no fault in the actor.⁷⁵ As fault liability rules condition liability upon the existence of the defendant’s fault, the defendant’s claim that his actions were justifiable or excusable, and hence that he was not at fault, is relevant to the ultimate question of liability. By contrast, strict liability rules do not condition liability upon the existence of defendant fault, and hence the defendant’s argument that he was not at fault does not affect the liability decision. In this case, asserting justifications or excuses is simply irrelevant.

B. THE ECONOMICS OF STRICT LIABILITY AND FAULT LIABILITY

The doctrinal Section explained the legal difference between strict liability and fault liability, but it did not explain why the law is structured this way. This Section uses economics to explain the function of the law and demonstrates how both strict liability and fault liability rules serve that function.

71. COLEMAN, *supra* note 18, at 224 (“tort liability is not generally defeasible by excuses.”); Joseph Raz, *Responsibility and the Negligence Standard*, 30 OX. J. LEGAL STUD. 1, 10 (2010) (“Excuses excuse from punishment and more, but are not relevant to compensation.”).

72. COLEMAN, *supra* note 18, at 220; JEFF MCMAHAN, *KILLING IN WAR* 44 (2009) (“Strict liability in tort law is liability that is defeasible neither by excuse nor by justification.”).

73. *Id.* at 217–20.

74. *Id.* at 217–18.

75. *Id.* Such defenses should also be distinguished from privileges or so-called “public policy defenses.” See generally GOUDKAMP, *supra* note 66, § 5.3.

1. *Economics Foundation*

Economic analysis rests on a consequentialist philosophical foundation that whether an action is right or wrong depends on whether its consequences are good or bad.⁷⁶ Whether conduct is good or bad depends on whether it creates greater benefits or costs for society. Conduct that creates greater benefits than costs is known as social welfare maximizing conduct.⁷⁷ This conduct may also be described as efficient behavior, as such actions allocate resources towards uses that yield optimal welfare results.⁷⁸ As humans usually try to act in ways that bring about greater benefits than costs, we often naturally act in welfare-maximizing ways.

However, in a subset of cases, people fail to act in welfare maximizing ways. This occurs because the private costs and benefits that an individual incurs from an action often differ from the social costs and benefits.⁷⁹ Most commonly, this happens when the cost of an individual's actions are borne not by himself, but by someone else. In such a case, the actor receives the benefit of his action but does not suffer the cost. This is known as a "negative externality." Since the actor receives greater benefit than he does cost, he will take the action. However, it may be that, when all of the benefit and cost for everyone in society is taken into account, the social cost of the action is higher than the social benefit. Accordingly, in this case the actor has an incentive to act in a way that reduces social welfare.

76. Specifically, it rests on a utilitarian basis. This article therefore departs from the view, once held by Richard Posner, that law and economics rests on a deontological (specifically Kantian) foundation. See Richard A. Posner, *Utilitarianism, Economics, and Legal Theory*, 8 J. LEGAL STUD. 103 (1979); Richard Posner, *The Ethical and Political Basis of the Efficiency Norm in Common Law Adjudication*, 8 HOFSTRA L. REV. 487 (1980). For a critique of such a position, see, for example, Jules L. Coleman, *Efficiency, Exchange, and Auction: Philosophic Aspects of the Economic Approach to Law*, 68 CALIF. L. REV. 221, 237–47 (1980); Jules L. Coleman, *Efficiency, Utility, and Wealth Maximization*, 8 HOFSTRA L. REV. 509, 525 (1980);

77. COOTER & ULEN, *supra* note 5, at 37–43; LOUIS KAPLOW & STEVEN SHAVELL, FAIRNESS VERSUS WELFARE 15–38 (2002). In discussing the maximization of welfare, this Article takes a different approach to law and economics than those that rest on maximization of wealth. See, e.g., Posner, *Utilitarianism*, *supra* note 76; Posner, *Efficiency*, *supra* note 76. This Article takes the view that wealth is at best a poor proxy for welfare. See generally Ronald Dworkin, *Is Wealth a Value?*, 9 J. LEGAL STUD. 191 (1980).

78. See RICHARD A. POSNER, THE ECONOMIC ANALYSIS OF LAW § 1.2 (2010). Accordingly, this paper uses efficiency in the sense of Kaldor-Hicks efficiency rather than Pareto Optimality. Resources are efficiently allocated in such a way that those better off could compensate those who are made worse off by the resource distribution.

79. COOTER & ULEN, *supra* note 78, at 39–40.

2. *The Economic Goal of Tort Law*

Tort law performs multiple functions. One aspect of the tort system is economic in nature. From an economic perspective, the function of tort law is to give people incentives to behave efficiently. In the absence of tort law, this would often not occur due to a negative externality problem. We can illustrate this problem with a hypothetical example.

Imagine that person A owns a house with a fireplace. There is a ten percent probability that a spark will escape and set fire to the roof of his neighbor's, B's, house. If this occurs, the damage to B's roof will be \$1000. Multiplying the amount of damage with the probability of its occurrence provides the expected accident costs—in this case \$100. Now imagine further that A has the option of buying a spark-catching device for a cost of \$80 that will completely prevent sparks from escaping. In such a scenario buying the device would maximize social welfare. Buying the device will impose a cost of \$80, but results in a benefit as \$100 in expected accident costs are forgone. The marginal cost imposed by the device is outweighed by the marginal benefit. However, while buying the device maximizes social welfare, it does not maximize A's private welfare. By not buying the device, A benefits by saving \$80 and the expected \$100 cost of this action is borne by B. Alternatively, buying the device would require him to pay \$80 and receive no benefit in return. Therefore, in the absence of legal regulation, A is unlikely to take the efficient action and buy the device.

On the other hand, in some cases the benefit of avoiding the harm would be outweighed by the cost of precaution. For example, imagine that the device costs \$110, not \$80. In this case buying the device would decrease social welfare: the cost of precaution outweighs the expected accident costs. Buying the device would impose a total cost of \$110 on society and only result in saving \$100. As the benefit is lower than the cost, the act of buying the device would be inefficient and therefore ought to be avoided, even though doing so may result in causing damage to B's roof.

The economic goal of tort law is to prevent externality problems like this one and provide individuals with incentives to behave efficiently. It accomplishes this goal through the imposition of liability. By making the actor pay a fee to the injured party (the externality bearer) the law shifts the costs of the action onto the actor. Doing so forces the actor to internalize the costs of his conduct.⁸⁰ Therefore, when deciding how to act, the actor's own

80. *Id.* at 190 (“The economic essence of tort law is its use of liability to internalize externalities created by high transaction costs.” (emphasis omitted)).

private cost-benefit analysis will take into account the full cost of his action. Thus, he will only act when the total benefit is greater than the total cost.

When imposing liability, tort law relies on two categories of liability rules: strict liability and fault liability. The next sections demonstrate how both strict liability and fault liability rules encourage the actor to behave efficiently. Before moving onto the precise workings of these rules, we must point out a definitional difficulty. Both doctrinalists and economists discuss strict liability rules and fault liability rules. However, when economists talk about strict liability, they typically mean harm-based strict liability.⁸¹ Likewise, when they discuss fault liability, they typically mean negligence rules. Economic literature contains little discussion of intentional fault in tort law; instead this is often covered in the discussion of criminal law.⁸² In keeping with this pattern, this Section shall discuss only the economics of harm-based strict liability and negligence rules.

3. *Strict Liability Rules*

Harm-based strict liability is liability imposed any time that a defendant's conduct causes a harmful outcome. In economic terms, this means that the actor will be liable every time he imposes a cost on someone else.⁸³ To see how such strict liability promotes efficient behavior on the part of the actor, consider the situation once again with A and B.

Imagine that the fire catching device costs \$80. In this situation, buying the device increases social welfare. Now, A has an incentive to act efficiently. When deciding whether to buy the device, A has two options: either buy the device for \$80, or do not buy the device and expect to pay \$100 in accident cost. Assuming that he is a rational welfare maximizer, A now has an incentive to buy the device. Doing so will result in him paying \$80 on the device rather than \$100 in expected liability.

Alternatively, if the device costs \$110, then the cost it produces is greater than the benefit. In such circumstances, A will not buy the device. Once again, he has two options: either buy the device for \$110, or do not buy the device and expect to pay \$100 in liability. As liability is the cheaper option, he has an incentive not to buy the device. Therefore, the operation of the strict liability rule creates incentives for the actor to behave efficiently.

81. *See, e.g., id.* at 201–04.

82. *Id.* at 188; MARTHA CHAMALLAS & JENNIFER B. WRIGGINS, *THE MEASURE OF INJURY: RACE, GENDER, AND TORT LAW* 64 (2010) (calling the economic analysis's theory of intentional torts as "precarious and marginal").

83. COOTER & ULEN, *supra* note 5, at 201–04.

4. *Negligence Rules*

As discussed in the preceding Section, under a negligence rule, the defendant will be liable only when he causes a harmful outcome through engaging in conduct that a reasonable person would not engage in. As seen, the law deems that a reasonable person would only take actions when the marginal benefit of the action outweighs the marginal cost. Thus, in the economic interpretation, the determination of a reasonable person is a question of whether the defendant behaved efficiently. Under a negligence rule, defendants will not be liable when their conduct is efficient, but will be liable when it is inefficient.⁸⁴

Consider the effects of this liability rule on the behavior of the hypothetical defendant, A. Firstly, consider the case where the device costs \$80. Once again, A has two options: buy or do not buy. He knows that if he does not buy the device and an accident results, the court will ask whether taking this risk was reasonable. As the cost of the device is outweighed by the benefit of buying the device, the court will find this unreasonable. In this scenario, A's expected liability is \$100. Alternatively, he could buy the device for \$80 and thus avoid liability completely. Therefore, he has an incentive to buy the device.

Alternatively, imagine the device costs \$110, and that buying it would decrease social welfare. Once again, A can either buy the device or not buy it. If he does not buy it and an accident occurs, he knows the court will ask whether this action was reasonable. As the cost of the precaution is greater than the expected cost of the accident, he knows that the court will deem his failure to take care to be reasonable. Therefore, if he does not buy the device, he spends no money on the device and pays no money in liability. As this is cheaper than buying the device for \$110, A has an incentive not to buy it, and once again acts in accordance with the demands of social welfare.

5. *The Substantive Difference between Strict Liability and Negligence*

Thus, both strict liability and negligence rules give the actor efficient incentives and promote social welfare.⁸⁵ Nevertheless, the rules achieve this

84. *Id.* at 205–08.

85. However, a simple negligence rule is potentially superior over a simple strict liability rule when one considers not only the rule's affect on the actor's incentives, but also the incentives of the victim. A negligence rule provides incentives for efficient levels of care from both parties, whereas strict liability does not. *See* COOTER & ULEN, *supra* note 5, at 204–08. This defect in strict liability can be overcome through the supplementary addition of a contributory negligence defense. *Id.* at 208–11. In which case, both negligence and strict liability rules can provide incentives for bilateral care. In such circumstances, which is the superior rule depends on how the rules affect the parties' activity levels. *Id.* at 211–13.

goal in diverging ways. Strict liability and negligence differ in how they distribute costs between the parties.⁸⁶ Strict liability holds the actor liable whenever his actions cause an accident, regardless of whether his actions are efficient. As the actor knows that he will be liable for every accident, the cost of his action is always internalized to him. And the person who initially bears the externality, the injured party, is not required to bear the accident cost. Compare this to the situation under a negligence rule. Now the actor is only liable when his actions are inefficient. If he acts efficiently, then he faces no liability. Therefore, he only internalizes the cost of inefficient behavior. When the actor does act efficiently, the externality bearer is the one who must bear the accident cost. Hence strict liability is more favorable for plaintiffs than for defendants.

III. STRICT LIABILITY AND FAULT LIABILITY IN COPYRIGHT INFRINGEMENT

The preceding Part demonstrated the doctrinal and economic differences between strict liability and fault liability rules. This Part will apply these insights to decide whether copyright infringement is a strict liability or fault-based tort. Section A summarizes the main doctrinal features of the copyright infringement action. Section B discusses existing theories of copyright infringement, highlighting the orthodox view that copyright infringement is a strict liability tort. Sections C and D explain the thesis that copyright infringement is a fault-based tort. Section C discusses this thesis from a doctrinal perspective while Section D uses a law and economics method. Finally, Section E responds to criticisms of the thesis and introduces one caveat into the analysis.

A. BASIC COPYRIGHT DOCTRINE

This Section introduces the reader to the *prima facie* case in a copyright action before discussing the most important affirmative defense, the fair use doctrine.

1. *The Prima Facie Case*

Upon fixing an original work of authorship in a tangible medium, the author automatically receives copyright protection over the work.⁸⁷ Under the 1976 Copyright Act, copyright protection provides the author with a

86. In addition to substantive differences, the two rules also differ in the level and type of administrative costs they create for the legal system. *See id.* at 223–24; RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 178–81 (2007).

87. 17 U.S.C. § 102 (2012).

bundle of exclusive rights.⁸⁸ Taken together, these rights provide the author with the exclusive ability to copy the work.⁸⁹ In order to establish a prima facie case of copyright infringement, the plaintiff must demonstrate two facts.⁹⁰ Firstly, he must show that the defendant copied from the plaintiff's work. Secondly, he must demonstrate that through this copying, the defendant produced a substantially similar work.

In order to prove copying, the plaintiff must demonstrate that the defendant either mechanically copied the work (e.g., by photocopying the work) or alternatively that the defendant had the plaintiff's work in mind when creating a new work.⁹¹ However, in this latter case, it is not necessary for the defendant to be consciously aware that he is copying from the plaintiff's work.⁹² This was most famously demonstrated in the *Harrisongs* case.⁹³ In 1971, former Beatle George Harrison was held liable for copying the Chiffon's hit single, *He's So Fine*, when creating his song, *My Sweet Lord*. Harrison argued that he did not consciously copy the song, and that if he did copy it, he did so without awareness of his actions. However, the court concluded that even subconscious copying constituted a copyright infringement. Harrison had heard the Chiffon's song in the past, and, when creating *My Sweet Lord*, subconsciously brought it to mind and copied its main elements. This was sufficient copying to impose liability.

Once copying is established, the plaintiff must show that this copying resulted in the production of a work that is substantially similar to the copyright holder's work.⁹⁴ In cases where the defendant has copied the work verbatim, this is an easy requirement to satisfy. In cases, such as the *Harrisongs* case, where the defendant has not copied verbatim, the plaintiff must prove that the audience for the plaintiff's work would perceive substantial similarities between the two works. Lastly, it is important to note that the plaintiff only satisfies the prima facie case if he can prove that the production of a substantially similar work was the result of actual copying.⁹⁵ If the defendant creates a substantially similar work, but did so through a

88. *Id.*

89. *Id.*

90. *Arnstein v. Porter* 154 F.2d 464 (2d Cir. 1946); *see also* 4 MELVILLE NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.03 (2010); 2 GOLDSTEIN, *supra* note 13, § 9.1.

91. 2 GOLDSTEIN, *supra* note 13, § 9.2.

92. *Id.*

93. *Bright Tunes Music Corp. v. Harrisongs Music, Ltd.*, 420 F. Supp. 177 (S.D.N.Y. 1976).

94. 2 GOLDSTEIN, *supra* note 13, § 9.3.

95. *See id.* § 9.2.

chance independent re-creation, then he is not liable no matter how similar the works are.⁹⁶

2. *Fair Use*

If the plaintiff proves copying and substantial similarity, he has successfully established a *prima facie* case against the defendant. The burden then shifts to the defendant to exculpate himself through the introduction of affirmative defenses. The most important of these defenses is the fair use doctrine.

According to the Copyright Act, the copyright holder's exclusive rights are granted "subject to" the fair use doctrine.⁹⁷ This doctrine establishes that it is "not an infringement" to copy a copyrighted work in cases where copying is "fair."⁹⁸ If the planned copying is a fair use, the copyst need not receive the author's consent in order to copy, nor pay a license fee, nor pay damages after the copying takes place. Since its inception, and eventual codification into statute, this doctrine has become a fundamental part of the copyright infringement analysis with application in a great variety of cases. For example the fair use doctrine has been applied to legitimize copying for the purposes of parody,⁹⁹ time-shifting television programs,¹⁰⁰ reproducing thumbnail versions of images,¹⁰¹ playing a political opponent's campaign theme music,¹⁰² digitizing books,¹⁰³ quoting from private and unpublished letters,¹⁰⁴ reverse engineering computer programs in order to create interoperable programs,¹⁰⁵ and displaying cached websites in search engine results.¹⁰⁶

Nevertheless, despite becoming one of the most venerated and important doctrines in copyright, it is also one of the most mysterious.¹⁰⁷ The term

96. *See, e.g.*, Calhoun v. Lillenas Publ'g, 298 F. 3d 1228 (11th Cir. 2002).

97. 17. U.S.C. § 106–07 (2012).

98. *Id.* § 107.

99. Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569 (1994).

100. Sony Corp. of Am. v. Universal City Studios, 464 U.S. 417 (1984).

101. Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003).

102. Keep Thomson Governor Comm. v. Citizens for Gallen Comm., 457 F. Supp. 957 (D.N.H. 1978).

103. Authors Guild, Inc. v. Hathitrust, 755 F.3d 87 (2d. Cir. 2014).

104. Wright v. Warner Books, Inc., 953 F.2d 731 (2d Cir. 1991).

105. Sega Enters. Ltd. v. Accolade, Inc. 977 F.2d 1510 (9th Cir. 1992).

106. Field v. Google, Inc., 412 F. Supp. 2d 1106 (D. Nev. 2006).

107. Professor Lawrence Lessig famously declared fair use so vague and unpredictable that it was nothing more than the "right to hire a lawyer." *See* LAWRENCE LESSIG, FREE CULTURE 187 (2004). However, more recently a number of empirical studies have sought to demonstrate that the doctrine is more consistent than first appreciated. *See, e.g.*, Barton Beebe, *An Empirical Analysis of U.S. Copyright Fair Use Opinions, 1978–2005*, 156 U. PA. L.

“fair” has no exact definition, and ultimately whether a use is fair is a question left for judicial determination. The breadth of the doctrine’s application coupled with its lack of succinct definition has resulted in commentators calling it “the most troublesome” doctrine in copyright law.¹⁰⁸

However, the Copyright Act does provide some guidance on the content and meaning of fairness. Firstly, it provides some illustrative examples of fair uses. According to the Act, copying is fair for the purposes of criticism, comment, news reporting, teaching, scholarship, or research.¹⁰⁹ Secondly, the Act provides a list of four non-exhaustive factors that courts ought to consider in determining whether a use is fair. Those factors are (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market or value of the copyrighted work.¹¹⁰ It has been said that this last factor, often known as the market harm factor, is “undoubtedly the single most important element.”¹¹¹

When discussing the four factors, two further points are relevant. Firstly, each factor is vague and leaves room for substantial judicial interpretation. Hence, courts discussing the purpose and character of the use have decided that whether a use is “transformative” (defined as altering the original work by adding “new expression, meaning, or message”) is an important consideration.¹¹² Likewise, when discussing the nature of the protected work, courts have drawn a distinction between fictional works (which receive greater protection) and factual works (which receive less protection).¹¹³ And secondly, because the factors are non-exhaustive, the court has room to

REV. 549 (2008); Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537 (2009); Neil W. Netanel, *Making Sense of Fair Use*, 15 LEWIS & CLARK L. REV. 751 (2011); Matthew Sag, *Predicting Fair Use*, 73 OHIO ST. L.J. 47 (2012).

108. *Dellar v. Samuel Goldwyn, Inc.*, 104 F.2d 661, 662 (2d Cir. 1939); *see also supra* note 107.

109. 17 U.S.C. § 107 (2012); although even for these enumerated uses, some may argue that they must also comply with the four factor analysis; *see, e.g.*, *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 561 (1985) (Congress “resisted pressures from special interest groups to create presumptive categories of fair use”); *Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1258 (2014).

110. *Id.*

111. *Harper & Row*, 471 U.S. at 566. However, courts have arguably retreated from this position in subsequent decisions. *See Sag, supra* note 108, at 63.

112. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994); *see also* Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105 (1994).

113. *See, e.g.*, *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. at 563; *Campbell*, 510 U.S. at 586.

supplement them with other considerations. Consequently, we see other considerations, independent of the four factors, affecting fair use determinations. One such consideration is whether the defendant acted in accordance with customary standards of fair dealing.¹¹⁴

Given the complex and open-ended nature of this doctrine, various scholars have appealed to the normative foundations of copyright in order to clarify when fair use does and ought to apply. In the remainder of this Section, we shall introduce two of the most important theories: the “market failure” theory and the “balancing of public interests” approach.¹¹⁵

a) The Market Failure Approach

In an influential article, Professor Wendy Gordon argued that fair use exists to cure market failures caused by copyright protection.¹¹⁶ Copyright protection attempts to cure a public-goods market failure: in the absence of exclusivity, expressive works will be under-produced. However, in some cases, the very existence of copyright protection will lead to further market failures. Most importantly, when transaction costs of licensing are high or when copying would lead to significant positive externalities for society, the market mechanism may break down. As the potential users cannot negotiate a license, society forgoes some socially beneficial copying. This suggests that fair use should exist to cure such copyright-induced market failures. That is, a use should be held fair when the planned use is socially beneficial and when the application of the normal copyright rights would result in a market failure which would prevent this socially beneficial copying from occurring.

The Supreme Court in *Sony Corp. of America v. Universal City Studios, Inc.*¹¹⁷ subsequently cited Professor Gordon’s article, and held that recording television programs for the purpose of watching them at a later time (i.e., time-shifting) is fair. One interpretation of this decision is that time-shifting

114. *Harper & Row*, 471 U.S. at 562; see Jennifer Rothman, *Copyright, Custom, and Lessons from the Common Law*, in SHYAMKRISHNA BALGANESH, *INTELLECTUAL PROPERTY AND THE COMMON LAW* 230, 237–40 (2013).

115. A third theory, a “harm-based” approach, proposed by Professor Christina Bohannon shall be discussed later in separating the concepts of harm and fault within copyright law. See Christina Bohannon, *Copyright Harm, Foreseeability, and Fair Use*, 85 WASH. U. L. REV. 969 (2007).

116. Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600 (1982). I have simplified Professor Gordon’s Analysis slightly for brevity purposes. Professor Gordon’s original analysis also required, as a third element, that the plaintiff would not be substantially harmed by the use. This present Article assumes that a work cannot be socially beneficial if it would involve “substantial” harm to the author’s economic interests.

117. 464 U.S. 417, 478 (1984).

is considered to be fair because it is socially beneficial, but unlikely to occur if subject to the normal copyright rules. As Professor Gordon argued, it is likely that the home users that engaged in such time-shifting would find the transaction costs of negotiating licenses for this use to be prohibitively high.¹¹⁸

b) The Balancing of Public Interests Approach

More recently, the market failure approach has come under scrutiny. In 2002, Professor Glynn Lunney has argued that the market failure theory of fair use was fundamentally flawed.¹¹⁹ In his understanding, this theory fails to adequately take into account the public goods nature of expressive works. Public goods are non-rivalrous, meaning that multiple people can use them at any given moment without depletion. From a static perspective, these uses are always welfare enhancing because they allow people to enjoy the work and obtain value from it without causing any cost to anyone else. In this case, creating legal exclusivity to prevent people from using the work always results in forgone welfare. Thus, we see the catch-22 in which copyright is situated. Without excludability, these works would not be created due to a public goods market failure, but when excludability is legally engineered, it necessarily results in another market failure where the good will be sub-optimally used. Ultimately, copyright always creates a market failure of some sort. Once it is acknowledged that copyright always results in market failure, the market failure concept “cannot serve as a useful guide” in determining when fair use ought to apply.¹²⁰

If we then remove the concept of market failure from the fair use analysis, the only criterion left to judge whether a use is fair is whether the copying would be socially desirable. Thus, Professor Lunney argues that fair use ought to apply when the copying would better serve the public interest. Whether the copying serves the public interest depends on a balance between two variables. On one hand, if the copying is allowed to continue without compensation, then this may “lead to fewer works of authorship by reducing the incentives to create such works.”¹²¹ On the other hand, allowing such copying may “improve the public’s ability to use, transform, or otherwise

118. Gordon, *supra* note 104, at 1653–57.

119. Glynn Lunney, Jr., *Fair Use and Market Failure: Sony Revisited*, 82 B.U. L. REV. 975, 993–96 (2002).

120. *Id.* at 996.

121. *Id.* at 977.

obtain access to the existing work.”¹²² This is known, infamously in copyright discourse, as the incentive-access paradigm.¹²³

Furthermore, Professor Lunney believes that courts already use the incentive-access paradigm as a guide to fair use determinations. In *Sony*, the Supreme Court argued that fair use exists to provide a “sensitive balancing of interests.”¹²⁴ According to the Court:

The fair use doctrine must strike a balance between the dual risks created by the copyright system: on the one hand, that depriving authors of their monopoly will reduce their incentive to create, and, on the other, that granting authors a complete monopoly will reduce the creative ability of others.¹²⁵

In Professor Lunney’s interpretation, the Supreme Court held that time shifting was fair use because such time shifting “yield[s] societal benefits” by “expand[ing] public access to freely broadcast television programs.”¹²⁶ While at the same time there was no evidence to believe that this time-shifting would cause significant market harm to the copyright holders. This reasoning reveals that the Court’s ultimate focus is simply upon the question of whether the copying was socially beneficial. Such statements indicating that a balancing test lies at the heart of fair use continue to today.¹²⁷

122. *Id.*

123. This article does not intend to examine in great detail the nature of the incentive-access paradigm. Striking the balance between these two variables has been called “the central problem in copyright law.” See William M. Landes & Richard A. Posner, *An Economic Analysis of Copyright Law*, 18 J. LEGAL STUD. 325, 326 (1989). As a result, enough literature already explores this issue. See particularly, Glynn S. Lunney, Jr., *Reexamining Copyright’s Incentive-Access Paradigm*, 49 VAND. L. REV. 483, 485–86 (1996); Oren Bracha & Talha Syed, *Beyond the Incentive-Access Paradigm? Product Differentiation and Copyright Revisited*, 92 TEX. L. REV. 1841 (2014).

124. *Sony Corp. of Am. v. Universal City Studios*, 464 U.S. 417, 456 n.40 (1984).

125. *Id.* at 454.

126. *Id.*

127. See, e.g., *Hustler Magazine, Inc. v. Moral Majority, Inc.*, 796 F.2d 1148, 1151 (9th Cir. 1986) (holding that the fair use doctrine “is a means of balancing the need to provide individuals with sufficient incentives to create public works with the public’s interest in the dissemination of information”); *Warner Bros. Entm’t Inc. v. RDR Books*, 575 F. Supp. 2d 513, 551 (S.D.N.Y. 2008) (“In striking the balance between the property rights of original authors and the freedom of expression of secondary authors, reference guides to works of literature should generally be encouraged by copyright law as they provide a benefit readers and students.” However, authors “should not be permitted to ‘plunder’ the works of original authors, without paying the customary price, lest original authors lose incentive to create new works that will also benefit the public interest.” (citations and quotation marks omitted)); *Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1273 (11th Cir. 2014) (noting that “fair use must operate as a ‘sensitive balancing of interests’” (quoting *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 584 (1994))); *Corporation of Gonzaga Univ. v.*

B. EXISTING THEORIES OF COPYRIGHT INFRINGEMENT

With an understanding of basic copyright doctrine, we are now in a position to ask whether the liability rule is strict or based on fault. This Section will summarize the existing scholarship on this question. We first introduce the orthodox view that copyright is strict and then turn to an alternative analysis provided by Professor Steven Hetcher.

1. *The Orthodox View*

Judges¹²⁸ and copyright scholars¹²⁹ routinely (if not ubiquitously) refer to copyright infringement as a strict liability tort.¹³⁰ Unfortunately for our

Pendleton Express, LLC, No. CV-14-0093-LRS, 2014 WL 4792032, at *8 (E.D. Wash. Sept. 25, 2014) (holding that fair use “is designed primarily to balance the exclusive rights of a copyright holder with the public’s interest in dissemination of information affecting areas of universal concern, such as art, science and industry” (quotation marks omitted)).

128. See, e.g., Shapiro, Bernstein & Co. v. H.L. Green Co., 316 F.2d 304, 308 (2d Cir. 1963) (noting the “harshness of the principle of strict liability in copyright law”); Religious Tech. Ctr. v. Netcom On-Line Comm’n Servs., 907 F. Supp. 1361, 1370 (N.D. Cal. 1995) (“Although copyright is a strict liability statute, there should still be some element of volition or causation”); Educ. Testing Serv. v. Simon, 95 F. Supp. 2d 1081, 1087 (C.D. Cal. 1999) (copyright infringement “is a strict liability tort”); King Records, Inc. v. Bennett, 438 F. Supp. 2d 812, 852 (M.D. Tenn. 2006); (“[A] general claim for copyright infringement is fundamentally one founded on strict liability.”); Gener-Villar v Adcom Grp, Inc, 509 F. Supp 2d 117, 124 (D.P.R. 2007) (“[T]he Copyright Act is a strict liability regime under which any infringer, whether innocent or intentional, is liable.”); Faulkner v. Nat’l Geographic Soc’y, 576 F. Supp. 2d 609, 613 (S.D.N.Y. 2008) (“Copyright infringement is a strict liability wrong in the sense that a plaintiff need not prove wrongful intent or culpability in order to prevail.”); Jacobs v. Memphis Convention and Visitors Bureau, 710 F. Supp. 2d 663, 678 (W.D. Tenn. 2010) (“Copyright infringement, however, is at its core a strict liability cause of action, and copyright law imposes liability even in the absence of an intent to infringe the rights of the copyright holder.”).

129. See, e.g., A. Samuel Oddi, *Contributory Copyright Infringement: The Tort and Technological Tensions*, 64 NOTRE DAME L. REV. 47, 52 (1989) (“Liability for direct infringement is imposed on a strict liability basis.”); Kelly Cassey Mullally, *Blocking Copyrights Revisited*, 37 COLUM. J.L. & ARTS, 57, 83 (2013) (criticizing copyright’s “harsh strict liability standard.”); Ben Depoorter & Robert Kirk Walker, *Copyright False Positives*, 89 NOTRE DAME L. REV. 319 (2013) (“[e]ven for affluent defendants, overcoming the Copyright Act’s strict liability standard is highly burdensome”).

130. The fact that copyright infringement is part of tort law is sometimes forgotten by both tort and copyright scholars. On a theoretical level, a tort is an actionable civil wrong. See GOLDBERG & ZIPURSKY, *supra* note 6, at 1. Tort law is the subject that defines and sanctions such wrongs. Copyright infringement is a wrong (i.e., an infringement of a right), see Balganesch, *Obligatory Structure*, *supra* note 22, that is actionable and occurs between private parties, and thus it is a tort and governed by tort principles in exactly the same way as battery, trespass, or nuisance are. Professor Hetcher has explained the relationship between copyright infringement and tort in some detail. Hetcher, *supra* note 15, at 1283–88; see also CANE, *supra* note 22, at 76. And, on a positive legal level, numerous courts have produced dictum recognizing the relationship between copyright and tort. See, e.g., Ted Browne Music

concerns, when making this claim, authors typically do not explain exactly what they mean by strict liability or why copyright fits into that category. We are often left guessing on this issue. Nevertheless, it appears that two factors are important in affecting the views of these authors. First, as part of the *prima facie* case, the plaintiff need not demonstrate how the defendant acted intentionally or negligently.¹³¹ The law apparently does not condition liability upon any of the common types of fault. And second, as evidenced by the *Harrisongs* case, copyright infringement can occur even when the defendant was unaware that he was copying from a previous work.¹³²

While this orthodox view is usually presented in an unsystematic fashion, one exception can be found in the work of Professor Shyamkrishna Balganesh.¹³³ Professor Balganesh provides a deeper analysis of the issue and concludes that copyright infringement is a form of harm-based strict liability. In order to be liable, the defendant must copy the protected work (conduct), this must lead to a substantially similar work (outcome), and this substantially similar work must cause market harm to the plaintiff (harm). Nevertheless, Balganesh argues copyright infringement is strange form of harm-based strict liability. The strangeness is that the issue of harm is not part of the *prima facie* case. Instead, the copyright holder need only show that the defendant copied (conduct) and this resulted in a substantially similar work (outcome), before the defendant is held responsible as a *prima facie* matter. After which, the defendant may raise the defense of fair use. Within the fair use analysis, the question of market harm is assessed largely under the fourth factor. As highlighted earlier, this market harm question was once viewed as “undoubtedly the single most important element.”¹³⁴ Thus, after the *prima facie* case has been established, the defendant may exculpate himself by demonstrating that his conduct did not cause the necessary harmful outcome

Co. v. Fowler, 290 F. 751, 754 (2d Cir. 1923) (“Courts have long recognized that infringement of a copyright is a tort.”); *Screen Gems-Columbia Music, Inc. v. Mark-Fi Records, Inc.*, 256 F. Supp. 399, 403 (S.D.N.Y. 1966) (“[Copyright] infringement constitutes a tort.”); *Screen Gems-Columbia Music, Inc. v. Mark-Fi Records, Inc.*, 256 F. Supp. 399, 403 (S.D.N.Y. 1966) (“[Copyright] infringement constitutes a tort.”). A more theoretically interesting question is whether copyright infringement is a singular tort or, because copyright actually grants multiple different exclusive rights, whether copyright infringement is in fact composed of multiple different unique copyright-based torts, see Patrick R. Goold, *Unbundling the Tort of Copyright Infringement* (unpublished manuscript on file with author)

131. See, e.g., Ciolino & Donelon, *supra* note 15, at 356 (noting that copyright infringement requires no “scienter, intent, knowledge negligence, or similar culpable mental state. On the contrary, liability for civil copyright infringement is strict”).

132. *Id.* at 352 (discussing the *Harrisongs* case).

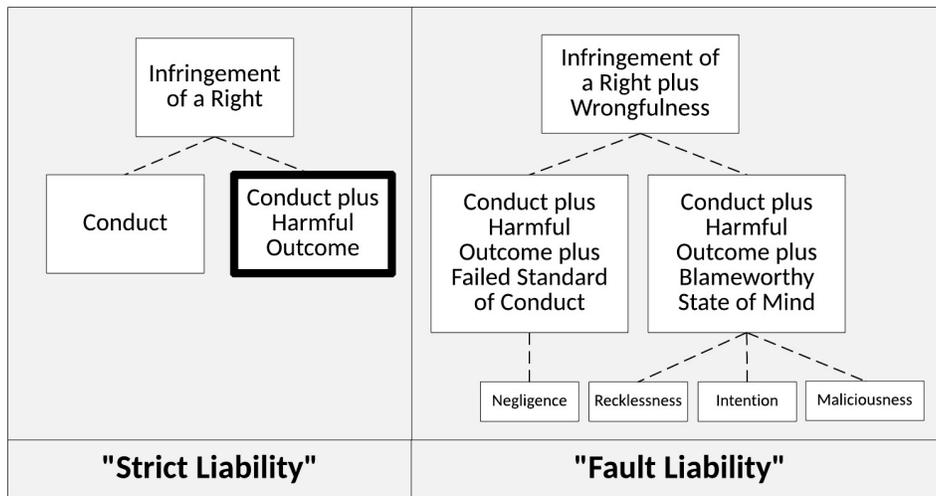
133. Balganesh, *supra* note 22, at 1682.

134. *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 566 (1985).

(i.e., harm to the plaintiff's market). Strangely, the law therefore relegates the question of harm to the defenses stage. Nevertheless, despite this strange treatment of the harm issue, fault is still not required. It "thus makes little difference for liability whether the copying was intentional, negligent, or a genuine mistake."¹³⁵

We can place this interpretation of copyright infringement as a strict liability tort on the diagram created earlier. Copyright infringement in the orthodox view is a harm-based strict liability tort and, as a result, the "conduct plus harmful outcome" box is emboldened to reflect that view.

Diagram 2



2. Professor Steven Hetcher's Fault Liability View

However, more recently one scholar has called upon us to revisit the question of whether copyright is indeed a strict liability tort. Professor Hetcher has argued that copyright infringement is in fact a fault-based tort.¹³⁶ This subsection firstly explains Professor Hetcher's argument before offering a critique.

135. Balganes, *supra* note 22, at 1682.

136. Steven Hetcher, *The Fault Liability Standard in Copyright*, in BALGANESH, *supra* note 114, at 431 [hereinafter Hetcher, *Fault*]; see also Steven Hetcher, *The Immorality of Strict Liability in Copyright*, 17 MARQ. INTELL. PROP. L. REV. 1 (2013); Wendy J. Gordon, *The Concept of "Harm" in Copyright*, in BALGANESH, *supra* note 114, at 452, 455 (confirming Professor Hetcher's position, with the addition that the notion of fault is also embedded in the doctrines of improper appropriation and substantial similarity).

a) Professor Hetcher's Argument

Professor Hetcher defines strict liability as liability imposed when a person causes “justiciable injury to the plaintiff,”¹³⁷ whereas fault liability rules requires the proof of an additional element other than causation: fault. Turning to copyright, Professor Hetcher then argues that proving causation is not enough to find infringement. A defendant is not merely liable because his copying caused the production of a substantially similar work, but it must also be the case that the copying “is not a legally recognized” fair use.¹³⁸ This question of fairness introduces a moral test that, adds something to the liability decision beyond the bare question of causation.¹³⁹

To clarify Professor Hetcher's position, it is helpful to compare it to that of Professor Balganesch. Balganesch argues that copyright is strict because liability is based upon conduct (copying), outcome (a substantially similar work), and harm (market harm).¹⁴⁰ When these elements are in place, a right has been infringed and liability follows; fault is irrelevant. By contrast, Professor Hetcher believes a right is infringed when the defendant engages in conduct (copying) and this causes a “justiciable” outcome (a substantially similar work), and thereafter the defendant can escape liability for this infringement by claiming that causing this harm was not wrongful (fair use). Thus, Professor Hetcher's analysis is distinct from Professor Balganesch's in two ways. Firstly, Professor Hetcher defines the concept of rights-infringement differently (and more broadly) than Professor Balganesch. Secondly, Professor Hetcher believes liability is conditioned on fault.

b) Critique

While there is much to admire in Professor Hetcher's conclusions, his reasoning is underdeveloped in places and therefore subject to attack. There are at least four important and interrelated critiques that Professor Hetcher's analysis does not completely address.

First, it is questionable whether a right is infringed when a defendant's copying causes the production of a substantially similar work. As noted earlier, autonomy rights are infringed through conduct alone, but other rights are only infringed when conduct causes harm. Professor Hetcher does not seem to believe that copyright is an autonomy right, because he acknowledges that causation is relevant to the liability decision. However, if he believes that copyright infringement adopts a harm-based strict liability

137. Hetcher, *Fault*, *supra* note 136, at 435.

138. *Id.*

139. *Id.*

140. *See supra* note 133.

rule, then he must explain why creating a substantially similar work is a harm sufficient to infringe the right. It is not clear what real world, non-legal harm is occasioned by the emergence of a substantially similar work.

Second, if unfair copying is a type of fault, then Professor Hetcher fails to explain its relationship to the other types of fault. If we consider Diagram 1 again, we know that Professor Hetcher would put copyright infringement on the right hand side of the diagram, but we do not know where it would be placed on that side. If copyright adopts a fault-liability rule, does that mean liability is based on negligence, intention, recklessness, or maliciousness? Or, does copyright adopt a completely unique fault standard? Professor Hetcher does not fully explain this point.

Third, the most natural reading of Professor Hetcher's article suggests that unfair copying is a unique type of fault (i.e., that unfair conduct is different from negligent, reckless, or intentional conduct). However, if unfair copying is not the same as preexisting categories of fault, why is it a fault standard in the first place? One cannot simply state that this doctrine is irrelevant to the question of rights-infringement, and is therefore a question of fault. Such reasoning would allow us to label as a fault inquiry any part of the tort action that is not relevant to the question of rights-infringement. Instead, one must provide a positive definition of the legal concept of fault, then demonstrate how this definition of fault covers the usual categories of wrongful conduct (i.e., negligence, recklessness, and intentionality), and then finally show how unfair copying is also covered by this definition.

The final critique builds on the preexisting two. Because the article presents no definition of legal fault, the fair use doctrine might not be a fault inquiry, but may be something else entirely. In particular, it may be a doctrine designed to determine whether a right infringement has occurred. Recall that in some interpretations of private nuisance, the reasonableness standard is said to bound the scope of the right, and therefore asking whether the defendant's conduct interferes with the right to reasonable use and enjoyment is simply a question of whether the defendant infringed the right. Likewise, the concept of fairness may in fact define the scope of the plaintiff's right. That is, the copyright holder's right may only be to prevent unfair uses of the work. Thus, when the court asks whether a use is fair, it may simply be attempting to determine whether a rights-infringement has occurred. If so, copyright liability is strict because liability is imposed any time the right is infringed. In order to be complete, Professor Hetcher's analysis needs to show why a fair use is a rights-infringement that is not wrongful, rather than conduct that simply does not infringe a right.

C. A DOCTRINAL REINTERPRETATION

Despite the problems with Professor Hetcher's analysis, this Article argues that his conclusion is correct: copyright infringement is a fault-based tort. Whether liability for copyright infringement is imposed upon the defendant depends upon the existence of four elements: conduct (copying), outcome (substantially similar work), harm (market harm), and fault (unfairness).

Nevertheless, the fault-based nature of copyright's liability rule frequently goes unnoticed because the formal structure of copyright infringement is a mess. After the plaintiff proves conduct and outcome, the inquiry moves to the question of fair use. Asking whether the defendant's copying was fair is a question of fault for the same reasons that asking whether a defendant's conduct was reasonable is a question of fault in a negligence action. But, unlike other fault-based torts, copyright infringement treats the concept of harm strangely. As noted by Professor Balganesch, the question of harm is not discussed as part of the *prima facie* case, but instead slips into the fair use analysis. Thus, what is ultimately a question of fault now takes on a dual character. The fair use analysis must determine whether the copying was harmful, and thus whether a rights-infringement has occurred, and thereafter, whether the copying was wrongful.

To illustrate this odd structure, this Section firstly explains how the question of fairness is a fault inquiry, much like the question of reasonableness in negligence. Secondly, it will demonstrate how the question of harm is bundled into the fair use fault inquiry. Doing so explicates the four elements of copyright infringement. Thirdly, this allows us to turn back to the orthodox view and diagnose the flaws in its reasoning.

1. *The Fault in Copyright Infringement*

Up until a point, Professor Balganesch and Professor Hetcher are in agreement. They both believe that copyright infringement is conditioned upon conduct (copying) and outcome (substantial similar work). Taking this as the starting point, the next question is whether copyright infringement is also conditioned upon fault.¹⁴¹ Is liability conditioned upon either the

141. Admittedly, by starting at this point, this Article must refrain from a detailed explanation of other doctrines in copyright that further refine the scope of the right, for example, the idea-expression doctrine, the first sale doctrine, and the list of additional limitations and exceptions appearing in §§ 108–122 of the Copyright Act. As these doctrines further refine the scope of the copyright, to engage in conduct covered by these doctrines is simply non-infringing. Therefore, to copy ideas or to sell a work after first sale is simply not a rights-infringement. Fair use has a different character, at least partially, because it justifies

defendant acting with a blameworthy state of mind or the defendant's failure to act in compliance with a standard of conduct?

a) A Blameworthy State of Mind?

Liability for copyright infringement is not conditioned upon the defendant's state of mind. The defendant need not copy intentionally, recklessly, fraudulently, with bad motives, or, as the *Harrisongs* case demonstrated, even consciously.¹⁴² Of course, saying that liability is not conditioned upon the defendant's mental state does not mean that state of mind is completely irrelevant in copyright law. Most obviously, it is taken into account at the remedies stage. The court has the discretion to impose statutory damages up to \$150,000 per infringed work when the infringement is "willful."¹⁴³ Importantly, however, these concerns are not relevant to the liability determination.

uses of certain works that are within the scope of the copyright holder's right on the grounds that such rights-infringement is simply not wrongful. My gratitude goes to Professor Rebecca Tushnet for pushing me towards greater clarity on this point.

142. One might argue that copyright infringement does take into account the copyist's mental state. Indeed, under the first fair use factor, the "purpose and character" of the defendant's use is important. The appeal to "purpose" does appear to make the defendant's subjective state of mind relevant. Additionally, some courts have, when assessing whether the use was transformative, asked whether the defendant had intended to communicate new meaning. *See, e.g.,* *Blanch v. Koons*, 467 F.3d 244, 259 (2d Cir. 2006) (quoting from Koons' affidavit explaining the defendant's artistic purpose behind the substantially similar use).

However, two points of clarification must be made in this regard. Firstly, even if mental state is relevant to the fair use determination, liability is not *conditioned* upon the defendant's mental state. While the relevant mental state may have some impact on the analysis, the absence of that mental state does not result in a finding of no liability. Unlike, for example, battery, there is no *requirement* that the defendant have a certain mental state before liability is imposed. The relevance of mental state to the fair use analysis is more akin to the relevance of mental state to negligent conduct: it might be some proof that the conduct was wrongful. If someone drives at 100 m.p.h. and is aware of that fact, his mental state may be taken into account when considering whether his driving was negligent. However, the tort of negligence is still based on standard-of-conduct fault because a certain mental state is not a requirement for liability. *See generally* Peter Handford, *Intentional Negligence: A Contradiction in Terms?*, 32 SYDNEY L. REV. 29 (2010).

Continuing this reasoning, one sees that the court's inquiry into the defendant's intent is something of a misdirection. Consider the transformative use cases. In such cases, the judge wishes to determine whether the use was transformative. Whether it is transformative or not has a bearing on the question of liability. To that end, the judge asks about the defendant's purpose when creating the use. He does so not because intent is legally relevant, but simply because the existence of an intent to communicate new meaning is some evidence that the use is transformative. Ultimately, the issue of whether the use is transformative, not the defendant's intent, has legal significance.

143. 17 U.S.C. § 504 (2012).

b) A Failure to Comply with a Standard of Conduct

Although liability is not conditioned upon a blameworthy mental state, liability is conditioned upon the defendant's failure to comply with a standard of conduct. The defendant in a copyright action is not liable merely because he copied a protected work leading to a substantially similar work; the copying must also be unfair. The fair use doctrine introduces a standard of conduct: fairness. It requires that when people copy, they do so fairly. It is only those who fail to reach this standard, and who copy unfairly, who will be held liable.

Therefore, copyright infringement is a fault-based tort in the same way that the tort of negligence is. Although these are very different causes of action, governing very different types of behavior, we see structural parity between them. In the tort of negligence, the defendant is not liable merely because he engages in proscribed conduct that causes an outcome (it is not enough that he takes a risk and this leads to some accident); it must also be the case that the conduct was unreasonable. Likewise, in copyright, it is not enough for the defendant to copy a work, leading to the production of a substantially similar work; it must also be the case that the copying is unfair.

Not only can we identify structural similarity between the tort of negligence and the tort of copyright infringement, but we can also see substantive similarity between the concepts of unreasonableness in negligence and unfairness in copyright infringement. As we refer to judging a defendant's conduct by a reasonableness standard as a "negligence rule," I suggest we refer to the practice of judging a defendant's conduct by a fairness standard as a "fairness rule." The remainder of this Section is dedicated to demonstrating the substantive relationship between these two liability rules.

i) The Relationship of the Negligence Rule and the Fairness Rule

It is greatly important to understand that the fairness rule and the negligence rule are not exactly the same type of liability rule. The fairness rule is not the negligence rule disguised in sheep's clothing. As a result, unfair conduct is not precisely the same type of fault as unreasonable conduct. Nevertheless, they belong in the same category of fault. The relationship of unfair conduct to unreasonable conduct is much like the relationship of intentional conduct to reckless conduct: they are part of the same category of fault, yet marginally different types of fault within the category. In both cases, the law sets a standard and sanctions those who fail to reach that standard. As a result, we see fundamental similarities between the concepts of reasonableness and fairness, three of which we will discuss here, before demonstrating the distinctions between these two types of fault.

(1) *Both reasonableness and fairness are defined in consequentialist terms.* As noted earlier, the reasonableness term has no precise or succinct meaning, but it is commonly defined in consequentialist terms. Whether conduct is reasonable depends on weighing the costs and benefits of the defendant's conduct. The only thing which changes in this regard are the types of costs and benefits that are salient to the calculation. In the tort of negligence, reasonableness requires a comparison of the costs of precaution and the benefit of avoiding the expected accident costs. Alternatively, in the tort of private nuisance, the reasonableness of the defendant's conduct requires comparing the gravity of nuisance and the social utility of the action.

Likewise, in copyright, fairness has no precise definition. But once again, whether conduct is fair seems to depend upon a cost-benefit analysis of the defendant's conduct. As Professor Lunney argued, whether copying is fair depends on the application of incentive-access tradeoff.¹⁴⁴ Copying is fair if the benefit that copying produces is greater than its cost in terms of reducing incentives for future authorial creation. On the other hand, copying is unfair if allowing it to continue absent liability will result in a reduction in future works and the value of this loss cannot be offset by benefits brought about by increased access. Therefore, if the copying has a net negative impact, then the copying will be considered unfair and the defendant will be liable to the plaintiff. The defendant must therefore pay damages for copying (or alternatively pay a license fee prior to the copying). On the other hand, if the copying has a net positive impact, then it will be considered fair and the defendant will not be liable.

Therefore, the incentive-access mantra of copyright plays the same role in guiding liability determinations as the Hand formula in the tort of negligence. In each case these formulas identify the variables that are relevant to the consequentialist balancing test, which in turn defines the relevant standard of conduct.

(2) *The wrongfulness in unreasonable and unfair conduct.* That fairness is defined in consequentialist terms also highlights the philosophical basis for labeling unfair copying as a type of fault. Both unreasonable conduct and unfair conduct are wrongful because they impose unwanted consequences on the rest of society. In the tort of negligence, unreasonable risk taking is wrongful because it exposes society to a greater than desirable number of harmful accidents. In copyright, unfair copying is wrongful because it deprives future society of new works that we find valuable. The future would seem much bleaker if it were to involve a vastly reduced number of expressive works.

144. See *supra* notes 122 & 123.

We can also view this feature from the opposite direction. Reasonable conduct is not faulty because it has good consequences for society: it saves resources by not spending them on needless precautions. Likewise, fair copying has good consequences for society because it results in greater access-benefits than incentive-costs.

(3) *Fault in the action, not the actor.* In both cases, therefore, fault occurs in the action, not in the actor. The individual who, with the best will in the world, negligently causes an accident has acted wrongfully because of the negative consequences he creates. The actor may not be at fault, but the action certainly is. Likewise, the defendant who unintentionally engages in unfair copying may not be personally blameworthy, but his actions are still wrongful because of their consequences. Thus, neither negligence nor copyright infringement depend upon the subjective mental state of the defendant. All that matters is the simple objective relationship between the conduct the defendant has performed and the standard society demands.

(4) *Distinguishing the fairness rule and the negligence rule.* Given the similarities, one may start to view the fairness rule as simply another version of the negligence rule. This is perhaps reinforced by the speed with which jurists turn to the concept of reasonableness when asked to explain the term “fairness” in a copyright context. The fair use doctrine has been called an “equitable rule of reason”¹⁴⁵ and as the ability to “use the copyrighted material in a reasonable manner.”¹⁴⁶ When Justice Story first laid the judicial foundations of the doctrine in *Folsom v. Marsh*, he described it as the freedom to use the copyrighted work for “fair and reasonable” purposes.¹⁴⁷

Nevertheless, fairness and reasonableness are different, and thus represent different types of fault. Elsewhere, Professor Gordon astutely pointed out that tort law is a law of harms.¹⁴⁸ Road traffic accidents, battery, private nuisance, defamation, all deal in the destruction or loss of something that society already has—whether that be physical health, use of property, or reputation. The situation is different in copyright. The social cost is not the destruction of something that society already possesses. Rather, the social cost is a foregone benefit. We enjoy expressive works, and hope they will be created in the future. Copying has the potential to reduce incentives, and thus result in future generations missing out on this benefit.

145. *Sony Corp. of Am. v. Universal City Studios*, 464 U.S. 417, 448 (1984).

146. HORACE G. BALL, *THE LAW OF COPYRIGHT AND LITERARY PROPERTY* 260 (1944).

147. *Folsom v. Marsh*, 9 F. Cas. 342, at 344 (C.C.D. Mass. 1841).

148. Wendy J. Gordon, *Of Harms and Benefits: Torts, Restitution, and Intellectual Property*, 21 J. LEGAL STUD. 449 (1992).

Against this backdrop we see the difference between reasonable conduct and fair conduct. Reasonable conduct is conduct that is simply the least harmful to society. Consider the hypothetical example of A and B once more. A could either expend a cost on buying the device to prevent an accident, or alternatively risk causing the accident. Neither situation is particularly desirable. It is simply the case that one choice is less bad than the other. While we term this conduct welfare maximizing, we may also think of it as incentivizing conduct that is simply the least welfare minimizing.

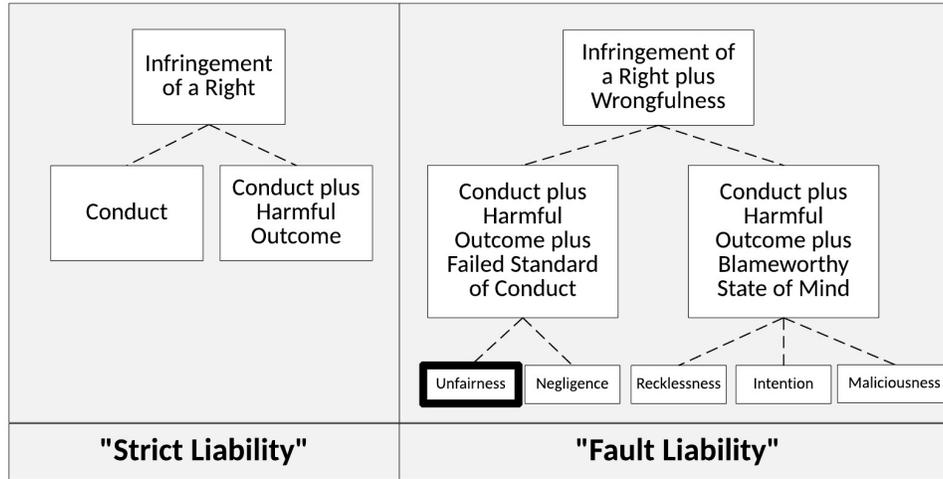
By contrast, fair conduct is much more positive. The copyist has two options: to refrain from copying or to copy. In cases where copying is unfair, and incentives are harmed, we mean that benefit is forgone, and society will lose out in the future. By contrast, when copying is fair, the benefit that is created in access is greater than the lost benefit resulting from reduced incentive. Unlike reasonable conduct, therefore, fair conduct is the maximization of benefit, not the minimization of loss. In engaging in benefit maximization, the defendant's conduct is rightful, and thus not subject to liability.

Thus unreasonable conduct and unfair conduct belong to the same class of fault—they both result in consequences that society would rather avoid. But society wishes to avoid unreasonable conduct because it results in unjustified detriment for society, while society wishes to prevent unfair conduct because it will result in the unjustified forgoing of benefit. As Professor Gordon highlighted, tort and copyright are “parallel mirror images” of one another; likewise so are unreasonable conduct and unfair conduct.¹⁴⁹

We can now attempt to summarize how copyright infringement fits on the diagram of liability rules. As Diagram 3 attempts to illustrate, unfair conduct is within the class of fault of failing to comply with a standard of conduct, but is not the same as negligent conduct. As a result, copyright infringement (represented again by the emboldening) falls in the fault liability section of the diagram, but is not associated with any of the recognized other liability rules. Instead, copyright infringement adopts the unfairness rule, which itself is unique to the world of copyright law.

149. Wendy J. Gordon, *Copyright as Tort Law's Mirror Image: "Harms," "Benefits," and the Uses and Limits of Analogy*, 34 MCGEORGE L. REV. 533, 533 (2002).

Diagram 3



2. *The Harm in Copyright Infringement*

So far, this Part has argued that liability for copyright infringement is based upon copying (conduct), substantial similarity (outcome), and unfairness (fault). The question now becomes, where does the concept of harm fit? Fault liability rules impose liability if the defendant wrongfully engages in the proscribed conduct that causes a harmful outcome. If the wrongfulness is the failure to comply with a standard of conduct, what is the harm in copyright infringement?

a) Isolating the Harm in Copyright Infringement

To answer the question of harm, we must think about the nature of the underlying right in copyright law. Once we understand the nature of the right, only then can we identify the type of activity that will interfere with the right.

According to the dominant understanding of copyright law in the Anglo-American world, the copyright is not an autonomy right, but a limited monopoly¹⁵⁰ to earn the market revenue generated by the work.¹⁵¹ It is not a

150. See, e.g., Samuelson, *supra* note 107, at 2617 (discussing the “limited monopoly conception of copyright”).

151. Not too long ago, scholars identified two different approaches to the scope of Anglo-American copyright. On one hand, there was the “minimalist” interpretation or the “incentives theory” approach. This approach viewed copyright (and patent) as a limited monopoly imposed to enable the author to receive a revenue for his work. The scope of the monopoly should be tailored narrowly so that the author should be able to recover his fixed

broad power to control all uses of the object. Instead, the right is an entitlement to satisfy the demand generated for the expression.¹⁵² As Justice

costs, and thus be endowed with the relevant incentives, but no further. On the other hand, there was a “maximalist” interpretation or “neoclassical” approach. This approach argued that the author ought to have a broad power of control over all uses of the work. Doing so would supposedly allow the author to internalize all the social value produced by the work. Therefore, not only would he have the incentive to create the work, but he would have the right *level* of incentive (i.e., an incentive level set by the market). This view based its theory largely on the Demsetzian theory of property that internalization via private rights will lead to socially efficient use of the good. See Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. 347 (1967). A number of scholarly works used this theory as grounds to interpret intellectual property rights broadly. See, e.g., Edmund Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265 (1977). On the difference between these two approaches, see, for example, Neil W. Netanel, *Copyright and a Democratic Civil Society*, 106 YALE L.J. 283, 308–11 (1996); Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 TEX. L. REV. 1031, 1042–43 (2005); Anne Barron, *Copyright Infringement, Free-Riding, and the Lifeworld*, in LIONEL BENTLY ET AL., *COPYRIGHT AND PIRACY: AN INTERDISCIPLINARY CRITIQUE* 93 (2010). However, despite the existence of these two approaches, the dominant view in the academic community today is that Anglo-American copyright is not a broad power of control, but a limited monopoly to earn enough revenue to cover fixed costs. This is in large part due to the work of a number of scholars that successfully critiqued the logic of the broad maximalist stance. See, e.g., Robert Merges & Richard Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 888 (arguing that a system of tailored incentives, not monolithic ownership, will best spur innovation); Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 TEX. L. REV. 1031, 1044–68 (2005); Mark A. Lemley, *Ex Ante Versus Ex Post Justifications for Intellectual Property*, 71 CHI. L. REV. 129 (2004); Mark A. Lemley & Brett Frischmann, *Spillovers*, 100 COLUM. L. REV. 257 (2007); Brett Frischmann, *Evaluating the Demsetzian Trend in Copyright Law*, 3 REV. L. & ECON. 649 (2007).

152. In this respect, copyright in the Anglo-American world is arguably slightly different from the authors’ rights systems of Continental Europe. Historically, many European jurists, basing their work on the philosophies of Kant in particular, viewed copyright not as a property right, but as a personal right. In this view, the author’s expression is not an object that is separable from the author, but is simply part of the author. In Kantian theory, the author’s speech is believed to be part of the author’s person (just as one’s organs, limbs, or thoughts are part of one’s person). And, as one has an innate right to control one’s person free from heteronomous interference, the author has a right to control how people interact with his or her speech. See generally Neil W. Netanel, *Copyright Alienability Restrictions and the Enhancement of Author Autonomy: A Normative Evaluation*, 24 RUTGERS L.J. 347, 363–82 (1993); ELEONORA ROASTI, *ORIGINALITY IN EU COPYRIGHT* 69–75 (2013). The result is that the author’s right is more of an autonomy right, in which the author has the exclusive ability to control what happens to the expression. This, of course, does not mean that the right is limitless and unbounded. Professor Drassinower’s work has demonstrated that a Kantian inspired right over expression in fact supports a robust public domain. See ABRAHAM DRASSINOWER, *WHAT’S WRONG WITH COPYING* (2015). However, it does mean that jurists must very carefully distinguish the author’s expression from other non-expressive concepts such as ideas, and design laws that respect the equal freedom of others in society to express themselves.

Breyer articulated in *Golan v. Holder*, copyright is the right to “charge a fee to those who wish to use the copyrighted work.”¹⁵³

As the right is to satisfy the demand generated by the work, it follows that the legally relevant harm occurs when the defendant’s actions interfere with the plaintiff’s ability to satisfy that demand.¹⁵⁴ Only when the plaintiff loses the market revenue has the defendant caused an outcome that impinges upon the right. Thus, in order to interfere with the right, the defendant must copy, leading to a substantially similar work that is capable of serving as a substitute for the plaintiff’s work, the existence of which diverts the market demand away from the plaintiff, resulting in lost sales and license fees.¹⁵⁵ In the absence of these elements, no rights infringement has occurred.

At this point, we can see that harm is also a necessary condition for copyright liability. As Professor Balganesch points out, harm is taken into account in the fair use analysis, particularly under the fourth factor. In the absence of market harm, the defendant’s use will be considered fair, and thus no liability will follow. This is most clearly demonstrated by the parody cases. In *Campbell v. Acuff-Rose*, the Supreme Court held that parodies are fair use. Even though parodies are often injurious to the copyright holder, the court held that parodies do not cause “cognizable market harm.”¹⁵⁶ As parodies merely suppress, and do not supplant, demand for the work, parodies belong to the class of cases where the defendant has copied (conduct) resulting in a substantially similar work (outcome), but there is no usurping of market

153. 132 S. Ct. 873, 899 (2012).

154. In this respect, this article’s concept of copyright harm is slightly broader than previous definitions. Professor Bohannon, *see* Bohannon, *supra* note 116, at 989, argues that harm in copyright can only occur when the defendant “usurps the copyright holder’s most foreseeable markets, or those markets which a reasonable copyright owner would have taken into account in deciding whether to create or distribute the copyrighted work.” While this definition of harm is plausible, this article believes that courts currently take into account harm to markets that were unforeseeable at the time of creation.

155. Professor Wendy Gordon has produced a more detailed analysis of the concept of harm in copyright. Approaching the issue from a philosophic perspective, Professor Gordon concludes that “copyright should recognize only the effects of rivalry as constituting ‘harm’” and licensing fees should only fall within the concept of copyright harm when the author has some “grave need” for them. Gordon, *supra* note 136, at 483. This Article agrees with Professor Gordon’s analysis. Departing slightly, however, it notes that currently courts appear to define the scope of the right to market revenue so that the defendant’s entitlement extends to “reasonable” licensing fees regardless of whether a “grave need” is present. *See* *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994) (holding that the test is whether the defendant’s use harms licensing market that are “traditional, reasonable, or likely to be developed.”). Therefore, market harm, as the law is currently formulated, extends not only to lost sales but also to some, but by no means all, lost licensing fees.

156. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 591 (1994).

revenue, and therefore no harm. As a result, cases such as these are instances in which no rights-infringement has taken place. These are “no harm” cases.

b) Harm and Fault in Fair Use

Isolating the role of harm in copyright infringement demonstrates a feature of copyright law that is often missed: the formal structure of this law is a mess. In other fault-based torts, whether the outcome was harmful and whether the outcome was wrongful are two separate questions. For example, imagine a battery where the assailant hits the plaintiff’s jaw. In relation to this outcome, there are two distinct questions: was the outcome harmful (can the defendant prove the contact either caused physical injury or was offensive enough to cause dignitary injury)? And thereafter, was the harm wrongful (was it brought about intentionally)? Or, imagine a tort of negligence case. Some accident has occurred, the victim then initiates a suit in which he is required to prove that the accident was harmful (under the doctrinal heading of damages), and thereafter that it was wrongful (by application of the negligence rule under the doctrinal heading of breach of duty).

The copyright infringement analysis currently bundles the questions of harm and fault into the same doctrine: fair use. On one hand, some fair use cases are best interpreted as “no harm” cases. In cases like *Campbell*, the defendant’s copying did not cause the legally relevant type of harm. On the other hand, there are “no fault” cases. As Professor Lunney demonstrated, the fundamental question rooted at the heart of the fair use analysis is whether the defendant’s copying was socially desirable. The question of whether the copying was fair is of the same character as the question of whether risk-taking is reasonable. If the conduct is beneficial for society, then, despite the harm it causes to the victim, the defendant’s actions are rightful and the defendant should not be held liable.

Interestingly, the demarcation between the “no harm” and “no fault” cases has become stronger since 2005. Prior to 2005, the fair use analysis was primarily concerned with the fourth factor issue of market harm. However, Professor Neil Netanel has demonstrated that post-2005, fair use analysis has undergone a paradigm shift.¹⁵⁷ Now, much intellectual heavy lifting is accomplished under the first factor. In particular, the issue that most prominently guides the outcome is the question of whether the use was “transformative.” If the use can be characterized as transformative, this will militate heavily in favor of a finding of fair use.

157. See generally Netanel, *supra* note 107.

The transformative works doctrine is important for our concerns because courts accept two different conceptions of transformation. On one hand, a use is transformative if it changes the expression. On the other hand, a use is transformative if it adds no new expression, but nevertheless adds new meaning. The first type of transformation is tied to the concept of “no harm.” In cases like *Campbell*, there was sufficient transformation of expression that the defendant’s copying did not lead to market substitution, and thus no legally recognizable harm. However, the latter type of transformation is related to the concept of “no fault.” In these cases, the courts often appeal to the social value produced by the copying. Although there is no new expression, and as a result there is potentially harm to the author’s market, courts are willing to justify these uses on the grounds that the transformed meaning creates benefits for society that outweigh that private harm. Three cases in particular demonstrate this point.

Perfect 10, Inc. v. Amazon.com, Inc. concerned the Google Image Search Engine and its reproduction of images.¹⁵⁸ Perfect 10 was an adult men’s magazine selling pornographic images of women. Google’s Image Search Engine searches the web for images, indexes those images and reproduces them in thumbnail form. Through this service, users were able to search for Perfect 10’s images. Google would then return to the user free thumbnail size versions of those photographs. Perfect 10 sought to obtain an injunction to prevent this.

The Ninth Circuit held that this was fair use. This conclusion followed despite the fact that there was arguable market harm. The trial court had found market harm existed because the reproduction of thumbnail images interfered with Perfect 10’s ability to sell thumbnail images in the cellphone market.¹⁵⁹ However, despite this fact, the Ninth Circuit held that Google Image Search provided clear social benefit in terms of increasing access to information and therefore was “highly transformative.”¹⁶⁰ As Professor Netanel notes, this is a case where Google’s display of thumbnail images was held fair because of the “use’s highly transformative, socially beneficial character despite possible harm to the plaintiff’s potential market for licensing thumbnails.”¹⁶¹

158. 508 F.3d 1146 (9th Cir. 2007).

159. *Perfect 10 v. Google, Inc.*, 413 F.Supp. 2d 828, at 851 (C.D. Cal. 2006) (“Google’s use of thumbnails likely does harm the potential market for the downloading of P10’s reduced-size images onto cell phones.”).

160. *Perfect 10*, *supra* note 158, at 1165.

161. Netanel, *supra* note 107, at 764.

A similar situation arose in *Bill Graham Archives v. Dorling Kindersley, Ltd.*¹⁶² That case involved concert posters of the rock band The Grateful Dead. The defendant was a book publisher who shrunk seven of the band's concert posters down to a reduced size to illustrate a coffee table book that chronicled the band's career. The Second Circuit held this to be fair use. The court found that the book used the images as "historical artifacts" and this use was important in "enhancing biographical information."¹⁶³ This was a distinctly different purpose from the previous use of the posters and therefore could be classified as transformative and thus fair. Interestingly, the court held that the transformation warranted a finding of no liability despite the strong argument that this case involved cognizable market harm. Bill Graham Archives was already in the market for licensing reduced size images. As a result, Professor Netanel argues that this case is quite striking because it demonstrates that as long as the use is transformative in meaning, "even actual market substitution is not enough to negate fair use."¹⁶⁴

Even more recently, the Google Books controversy also appears to fit into the "no fault" category.¹⁶⁵ Google copies and digitizes copyright protected books, and then allows users to access "snippet" views of the books. There is an arguable case that allowing users to access snippet views causes market harm.¹⁶⁶ As the user can access parts of the defendant's expression via Google, it is likely that some will use Google's search engine to access parts of the book rather than buy their own copy. However, the court found the use was transformative. Tellingly, the court held:

Google's use of the copyrighted works is highly transformative. Google Books digitizes books and transforms expressive text into a comprehensive word index that helps readers, scholars, researchers, and others find books. Google Books has become an important tool for libraries and librarians and cite-checkers as it helps to identify and find books. . . . Similarly, Google Books is also transformative in the sense that it has transformed book text into data for purposes of substantive research, including data mining and text mining in new areas, thereby opening up new fields of research.¹⁶⁷

162. 448 F.3d 605 (2d Cir. 2006).

163. *Id.* at 610

164. Netanel, *supra* note 107, at 760.

165. Authors Guild, Inc. v. Google Inc., 954 F. Supp. 2d 282 (S.D.N.Y. 2013).

166. Putting aside momentarily whether the negative impact is outweighed by other beneficial effects, on which see *infra* note 224.

167. *Id.* at 291.

Later the court characterized these functions of the service as providing “significant public benefits.”¹⁶⁸

This author interprets these three cases, and others,¹⁶⁹ as ones in which the plaintiff has suffered market harm. The defendant has copied the expression in way that arguably leads to market substitution. Nevertheless, this harm is justifiable. In each case, the copying was associated with strong public benefits. Thus, in each instance, the copying was fair because it created greater access-benefits than incentive-costs.

Finally, the realization that fair use bundles these two inquiries together solves a longstanding problem in copyright law. For decades, a debate has surrounded the nature of fair use. On one side, some believe that a fair use is an infringement of a right that we merely allow.¹⁷⁰ On the other hand, some believe that fair use is simply not an infringement at all. This Subsection demonstrates that both sides of the argument are half-correct. Some fair uses are outside the scope of the right. The no-harm cases like *Campbell* involve no recognizable harm, and are therefore not a rights-infringement. On the other hand, the three cases just discussed are cases where harm is quite likely, and therefore involve rights-infringement, but we allow this infringement because the conduct is good/not wrongful.¹⁷¹

168. *Id.* at 293.

169. *See* *Williams & Wilkins v. United States*, 487 F.2d 1345 (1973) (holding that a library’s wholesale copying of journal articles was fair use).

170. *See, e.g.*, ALAN LATMAN, STAFF OF S. COMM. ON THE JUDICIARY, 86TH CONG., COPYRIGHT LAW REVISIONS: STUDIES PREPARED FOR THE SUBCOMMITTEE ON PATENTS, TRADEMARKS, AND COPYRIGHTS OF THE COMMITTEE ON THE JUDICIARY, STUDIES 14–16, at 6 (Comm. Print 1960) (“Fair use may be viewed from two standpoints. It may be considered a technical infringement which is nevertheless excused. On the other hand, it may be deemed a use falling outside the orbit of copyright protection and hence never an infringement at all.”). Recognizing that some fair copying is simply good copying outside of the scope of copyright also pushes against the view that fair use will decrease in the future as potential licensing options increase. An alternative to viewing the fair use doctrine as primarily a fault doctrine is to view it as mere tolerated use. The theory goes that we tolerate such uncompensated use internalizing the external benefit to the author would be too costly. However, if this is the true nature of fair use, then as licensing becomes cheaper through the advent of better communications technology, then surely fair use will begin to shrink. *See generally* Rebecca Tushnet, *All of This has Happened Before and All of This Will Happen Again: Innovation in Copyright Licensing*, 29 BERKELEY TECH. L.J. 1447 (2014). Alternatively, if one views fair use as a fault standard which excuses those whose copying is socially beneficial, this shrinking seems less likely. It is not immediately apparent how improved licensing technology would change how a use would affect the incentive and access variables on which the fault determination rests.

171. One might also speculate as to the reasons why this bifurcation of fair use has become stronger post 2005. This Article offers a very tentative explanation. It is believed that tort law adopted fault liability in the mid-nineteenth century in order to meet the

3. *Responding to the Orthodox View*

Thus, copyright infringement is a fault-based tort. Liability is imposed on the basis of four elements: copying (conduct), substantial similarity (outcome), market harm (harm), and unfairness (fault). However, this characterization is often misunderstood because the harm inquiry is bundled into the fault inquiry. Nonetheless, with this knowledge, we are now in a position to evaluate some of the potential errors underlying the orthodox view that copyright is a strict liability tort.

Firstly, the orthodox view bases its belief that copyright infringement is strict on the premise that copyright law makes no mention of the usual types of fault. This is undoubtedly true; liability in copyright is not conditioned upon intentionality or negligence, nor recklessness, fraudulence, or maliciousness for that matter. It is nevertheless conditioned upon unfairness. Unfair conduct is not the same as intentional conduct or negligent conduct, but it is nonetheless a type of fault. The orthodox view has arrived at the conclusion that liability in copyright is strict through committing two errors. Its analysis begins by forgetting that failing to comply with a standard of conduct is a class of fault.¹⁷² Thereafter, it fails to see that unfair conduct is just as much the failure to comply with a standard of conduct as unreasonable conduct. Yet, given that copyright law deals with a different scenario to most torts—the regulation of future benefits rather than the regulation of detriment—it is hardly outlandish to believe that that the law has developed a unique type of fault to judge defendants' conduct.

challenges of the industrial revolution. See HORWITZ, *supra* note 1, 97–101. Similarly, one could argue that in the twenty-first century, the global economy has become far more digital. It is plausible that, as informational goods become increasingly important to the economy, it no longer is efficient to hold all those who cause harm liable, and that, instead, before a court holds a particular defendant liable, it must first engage in a consequentialist balancing test, much as they do in negligence. Therefore, as tort adopted negligence in response to the changing economic conditions of the industrial age, likewise, copyright has adopted a form of fault liability rule in response to the challenges brought by the digital revolution.

172. Various judges and scholars incorrectly associate the concept of fault only with a blameworthy state of mind. See, e.g., Ciolino & Donelon, *supra* note 15, at 356 (copyright infringement requires no “scienter, intent, knowledge negligence, or similar culpable mental state”); *Jacobs v. Memphis Convention and Visitors Bureau*, 710 F. Supp. 2d 663, 678 (W.D. Tenn. 2010) (“Copyright infringement, however, is at its core a strict liability cause of action, and copyright law imposes liability even in the absence of an intent to infringe the rights of the copyright holder.”). It is interesting to note that in patent law, there is at least one article pointing out how the legal concept of fault is broader than the notion of culpable mental state; see Jason A. Rantanen, *An Objective View of Fault in Patent Infringement*, 60 AM. U. L. REV. 1575 (2011).

Secondly, the orthodox view argues that copyright must be a form of strict liability because it holds subconscious copiers liable. But once again, this argument rests on an unduly narrow view of fault. This view stems from the incorrect assumption that the only fault recognized in tort is acting with a blameworthy state of mind. As copyright holds liable those, like subconscious copiers, who act without a blameworthy state of mind, it must be a form of strict liability. However, as seen, fault may be established where the defendant simply fails to comply with a standard of conduct. In these cases, even if the defendant acts with the best possible state of mind or the complete absence of a state of mind, he is nevertheless at fault if his actions fail to comply with the standard. Hence many people are liable in negligence even though they do not act with a bad state of mind. Likewise, in copyright infringement, the copying must still be unfair, and therefore wrongful, before liability will be imposed. Even if the defendant copies subconsciously, he has performed conduct that is wrongful and will accordingly be held liable for the consequences of that action.

Interestingly, there is a better argument that the orthodox view could make in relation to the subconscious copying rule. Proponents of this view could argue that, because of the subconscious copying rule, defendants are held liable even when their conduct is not volitional. The argument would rest on the assumption that subconsciously produced conduct is not voluntary conduct. Nevertheless, while this assumption would seem plausible, the argument that is based upon it would still fail.

If subconscious copying is not volitional, copyright would not be a form of strict liability, as strict liability requires that conduct be volitional. If copyright were to ground liability on the basis of involuntary conduct, it would impose a form of ultra strict liability on defendants. This would breach basic notions of responsibility in tort law.

Moreover, while there is an arguable case that subconscious copying is not volitional, tort law as it currently stands does not take the same view. Subconscious activity may not be volitional for purposes of criminal punishment, but it is volitional for purposes of civil liability. This is most clearly demonstrated by examining how tort law approaches the issue of so-called “automatic” actions. As Professor Peter Cane describes:

An experienced driver, for example, will do many things automatically or “without thinking” or “inadvertently” which a learner would do deliberately and attentively. The crucial difference between involuntary acts and automatic acts is that the former are uncontrollable whereas the later are controllable but not

consciously controlled. Far from being exempt from tort liability, automatic behavior is frequently the very essence of tortiously negligent conduct.¹⁷³

Automatic actions are a form of subconscious conduct. Such actions are not consciously brought about, but they are nevertheless technically within the defendant's sphere of control. This is sufficient volition for liability in tort law. Likewise, subconscious copying is volitional enough to find liability in copyright.

D. AN ECONOMIC REINTERPRETATION

The thesis that copyright infringement adopts a fault liability rule also gains support from economic theory. As this Section shows, the fairness liability rule distributes the cost of inefficient conduct from the externality bearer onto the actor in the same fashion as negligence rules.

1. *The Economic Goal of Copyright Law*

Like tort, copyright serves various goals. One of which is economic in nature. From an economic perspective, the function of tort and copyright is the same: to give people incentives to take efficient, welfare-maximizing action. In copyright, whether copying is welfare maximizing depends on the comparison of the effect of the copying on the incentive and access variables.

As in tort law, often individuals do not take welfare-maximizing action when dealing with copyrighted works because of a negative externality. When an individual unfairly copies, he forces a cost on future society—as authorial incentives will drop fewer works will be created. This cost may be greater than any benefit copying produces in terms of greater access. In such cases, copying is inefficient. However, from a private perspective, the copyist still has an incentive to engage in this behavior. Assuming that the duplication process requires no resources, then copying results in no cost to him personally, and only results in his benefit. The copyist discounts the cost this causes in terms of lost incentive to create because this cost does not fall primarily on him, but on future society.

As in tort law, copyright law uses liability to solve this problem. By holding the defendant liable for copying, we force him to internalize the costs of his conduct, and thus give him an incentive to behave efficiently. The question is, what type of liability rule is used to accomplish that goal?

173. CANE, *supra* note 22, at 30.

2. *A Strict Liability Rule in Copyright?*

We saw earlier that, in the economic understanding, strict liability is liability imposed anytime the defendant inflicts a cost on someone else, regardless of whether creating that cost was efficient. In copyright, we have defined cost as the lost incentive for authors to create. However, this is a difficult variable to quantify. Such a determination requires the court to consider how the defendant's copying will affect the actions of a group of people who are not present before it and who all have diverse motivations for creating.

Given the difficulty of assessing this variable, the court uses a proxy in valuing the lost incentives: harm to the plaintiff's market.¹⁷⁴ From a static point of view, harm to the author's market is irrelevant. If market harm has occurred, this is a sunk cost. Imposing liability does not eradicate it in any way. Furthermore, the work in question is already created; harming the author's market mercifully does not change that fact. Statically, therefore, liability is simply an expensive redistribution of the cost from one party to another. The only way such redistribution can be justified is when it is viewed from a dynamic perspective. If the defendant's copying causes this author market harm, then allowing it to continue will likely cause market harm to authors in the future. Future authors will perceive this situation negatively and their incentive to create will consequently decrease.

If copyright adopted a strict liability rule and imposed liability every time the defendant's actions caused a cost, liability would be imposed any time the copying negatively affected future authorial incentives. Since market harm is the proxy for lost future incentives, this would result in liability every time the copying causes market harm. In every instance where the defendant's actions threaten to produce lost sales, the defendant would be required either to negotiate a license *ex ante* or pay a damage award *ex post*. Thus, the harm to the author's market would be internalized to the copyist and accordingly he would take the cost to authorial incentives into account when deciding whether to copy.

However, copyright does not impose liability on every defendant who causes the copyright holder market harm. Under the balancing approach to fair use, the defendant's suffering of market harm is not sufficient to find liability. As Professor Lunney elaborates, the existence of market harm is

174. Lunney, *supra* note 119, at 1014. Professor Lunney highlights that the author's market harm is merely a proxy for lost future incentives but also notes that it is far from a perfect proxy. Ideally, the defendant who suffers harm should also demonstrate how that harm is likely to translate into lost incentives for future creativity.

simply the first stage of the fair use analysis. The ultimate and final arbiter in the fair use analysis is not whether the defendant was harmed, but whether the defendant's conduct will lead to greater access-benefits or incentive-costs. Therefore, copyright does not always impose liability on the defendant for the creation of social cost. Instead liability is only imposed on defendants in relation to their inefficiently caused costs. Causation is not the touchstone of copyright liability, but instead liability follows an analysis of "whether, on balance, society would be better or worse off by allowing the use to continue."¹⁷⁵ Even in cases where copying probably "decreases revenues to some extent," such as private copying via file-sharing networks, copying "may nevertheless expand access to an existing work substantially more for any given reduction in revenue."¹⁷⁶

3. *A "Negligence Rule" in Copyright*

Unlike the strict liability rules, negligence rules hold defendants liable only when the conduct was unreasonable. In the economic interpretation, this means liability is imposed only when the defendant's conduct is inefficient. Whether the conduct is inefficient depends on whether it produces greater benefits or costs.

Copyright infringement holds defendants liable when the copying is unfair. If the balancing interpretation of fair use is correct, then whether copying is fair depends on a balance of the cost to incentives versus benefits of increased access. Thus, whether copying is fair is a question of efficiency and welfare. Copyright has thus the same economic characteristics of a negligence rule in that a defendant's copying will only attract liability when it is inefficient.

As a result, we see that the liability regime in copyright distributes costs in the same way that a negligence regime does. We noted that, unlike a strict liability regime where the defendant always internalizes the total cost of his action, the defendant judged by a negligence rule only internalizes the cost of his inefficient action. In these cases, the cost of efficient action is borne by the externality bearer. In copyright, when deciding whether to copy, the defendant knows that liability will only result if his actions are inefficient. Thus, he only internalizes the harm to the author's market when the cost to incentives that such harm represents is not outweighed by the benefit of greater access. If his actions are efficient, the cost this causes in terms of lost incentive remains on the externality bearer: future society.

175. *Id.* at 1023.

176. *Id.* at 1026.

4. *Incentives in Copyright: Strict Liability or Fairness?*

Recall that Part I explained that both strict liability and negligence rules provide the defendant with an incentive to act efficiently. However, there is an important reason why this is not the case in copyright law. In the copyright context, adoption of strict liability would lead to inefficient action.

In tort law, the defendant usually captures the total benefit of his action, while the cost is typically borne by others. Strict liability holds the defendant liable for the cost he creates. The defendant's private cost-benefit analysis then accurately reflects a comparison of the total societal cost and total societal benefit. Thus, the defendant has an incentive to act in accordance with the demands of social welfare.

In some cases, applying a strict liability regime in copyright would have the same effect. Providing that the benefit of copying falls entirely upon the copyist, holding the defendant liable for all the market harm he creates will result in the defendant taking into account the entire social cost of his copying and the entire social benefit of his creation. The result of his cost-benefit analysis under such conditions will be efficient action.

However, the typical tort reasoning does not apply when the defendant's action creates not only negative externalities but positive externalities. If the defendant's action not only imposes costs upon one group of people, but also bestows benefits on another group, then strict liability will be inefficient. Such liability will result in the defendant taking into account the total social cost of his action, but not the total social benefit of his action. This raises the possibility that a type of action will be efficient—its total social benefit will outweigh its cost—but the defendant will fail to take it because the social/private cost outweighs the private benefit.

This is a problem that copyright law faces. While in some cases the total benefit of copying will fall on the defendant, in many cases the copying will result in positive externalities.¹⁷⁷ I noted previously that the benefit of copying is the benefit of access. Typically these access benefits fall largely upon people other than the copyist.¹⁷⁸ For example, when the defendant's copying results in a more competitively priced work, typically a large amount of the benefit is captured by the consumers who receive the work at a cheaper cost. Or, consider the case where the defendant copies to create a new work. It is unlikely that the defendant in such cases will be able to capture the entire positive value he creates in such an action, and much of the benefit will therefore remain with people who enjoy the new work. In

177. As is well known to copyright scholars. See, e.g., Gordon, *supra* note 116, at 1630–32.

178. See generally Lemley & Frischmann, *supra* note 151.

such cases, adopting a strict liability rule would be socially harmful. It would result in the copyist's own cost-benefit analysis taking into account the total social cost of his action but not the total social benefit. Thus, it would be likely that some copying that is beneficial would not occur.

The fairness rule that copyright adopts, with its economic characteristics of a negligence rule, avoids this problem. If the copying is efficient and produces greater benefits in the access than harm to incentive, then the defendant does not internalize the cost. Therefore, his own private cost-benefit analysis takes into account whatever private benefit he receives from copying and none of the cost. Assuming that he faces no cost involved in actually making the copy, his private benefit outweighs the private cost, resulting in efficient action. Alternatively, if the defendant's copying results in greater harm to incentives than to access, the defendant will have no incentive to engage in this behavior. In this case, the defendant will be liable, forcing him to internalize the cost he creates. Thus, his own private analysis reflects a balancing of the total social cost against his own private benefit. As social cost is greater than social benefit, in such cases social cost will necessarily be higher than the copyist's private benefit. Therefore the defendant will have an incentive to refrain from copying. Accordingly, whereas strict liability may lead to inefficient action, the fairness rule adopted by copyright leads to efficient action.

E. CRITIQUES, COUNTER-ARGUMENTS AND A CAVEAT

So far, the thesis presented is that copyright is a fault-based tort because liability hinges upon four elements: conduct (copying), outcome (substantially similar work), harm (market harm) and fault (unfairness). As the type of fault is a failure to comply with a standard of conduct, we see how copyright infringement distributes costs in a similar manner to negligence rules, thus leading to the copyist internalizing the negative externalities of his conduct and thereafter behaving efficiently.

However, this thesis may be critiqued in two important ways. Firstly, one may claim that fair copying is simply conduct that does not fall within the scope of the right. Secondly, some may believe that the position of fair use as an affirmative defense means ultimately that the underlying liability rule is still strict. This Section responds to each of these critiques in turn. Finally, one important caveat is highlighted.

1. *Fair Copying Is Outside the Scope of the Right*

The first critique is identical to the fourth question posed to Professor Hetcher earlier: is fair copying technically an infringement of the right that is justified because it is not wrongful, or is fair copying simply outside the

scope of the copyright, and thus not an infringement at all? So far, this article has argued that it is the former. However, one may argue that it is the latter. If this critique is correct, then copyright conditions liability solely upon rights-infringement, and is therefore strict.

a) Merit to the Critique

There is significant merit to this critique. Evidence in support of this argument can be found in two sources. Firstly, the statutory wording of the Copyright Act arguably substantiates this claim. The Copyright Act says that the copyright holder's exclusive rights are "subject to" fair use and that fair copying is "not an infringement."¹⁷⁹ Furthermore, the act says explicitly that "[a]nyone who violates any of the exclusive rights of the copyright owner . . . is an infringer of the copyright"¹⁸⁰ and that infringers of rights are liable for damages.¹⁸¹ Taken together, these statements provide some support for the view that (a) liability is imposed solely upon the infringement of a right, and (b) the concept of fairness qualifies the legal right, meaning that a fair use is simply not an infringement.¹⁸²

Secondly, the history of copyright law in the United States suggests that fair use was originally conceived of as conduct outside the scope of the entitlement. Initially the right of the copyright holder was a simple right to prevent near verbatim copying. In 1841, this changed with the case of *Folsom v. Marsh*.¹⁸³ As Professor Oren Bracha explains, this case typically viewed as the origin of fair use, but ironically the decision actually expanded the scope of the copyright right. Whereas before, the right was simply to prevent near verbatim copies, this case altered the scope of the right so that any copying would be labeled as an infringement if it interfered with the market for the plaintiff's work. The concept of "fairness" that Justice Story introduced simply represented those instances of copying that did not interfere with the copyright holder's market and thus did not fall within the scope of the right.¹⁸⁴ However, the fact that fair use began its life as conduct outside the scope of the right does not mean this is the situation today. It is entirely

179. 17 U.S.C. § 106–107 (2012).

180. *Id.* § 501(a).

181. *Id.* § 504(a).

182. See also Matthew Sag, *Copyright and Copy-Reliant Technology*, 103 NW. U. L. REV. 1607, 1609 (2009) ("[T]he fair use doctrine renders certain otherwise infringing actions relating to copyrighted works noninfringing.").

183. 9 F. Cas. 342 (1841).

184. Oren Bracha, *The Ideology of Authorship Revisited: Authors, Markets, and Liberal Values in Early American Copyright*, 118 YALE L.J. 186, 229 (2008); see also Matthew Sag, *Prehistory of Fair Use*, 76 BROOK. L. REV. 1371 (2011).

plausible that copyright has evolved over time and new meaning has been attributed to the fair use doctrine.

b) Counter-Argument to the Critique

To respond to this critique, we must distinguish between two concepts: on one hand, a standard that is employed to limit the scope of a plaintiff's right, and, on the other, a standard that distinguishes rightful from wrongful infringements of that right. Distinguishing these concepts is not easy and has gone under-theorized in tort law. However, we can find some answer to this question by looking into the literature on private nuisance.

As noted earlier, private nuisance contains the concept of reasonableness. But reasonableness has potentially a dual meaning. On one hand, the court could ask whether the defendant infringed the right to reasonable use and enjoyment, and on the other, the question could be whether the defendant's conduct was reasonable. As highlighted by Professor Gregory Keating and Professor Richard Wright, the difference between these two inquiries is a difference between impact and conduct.¹⁸⁵ In the first case, the question focuses solely on the impact to the plaintiff. Was the plaintiff harmed in a legally cognizable way? This is an impact analysis and the legal standard's role is to define the scope of impact that will be actionable. By contrast, the second case does not focus on the plaintiff but the defendant. The question is whether the defendant's actions were good or bad. This is not an impact analysis, but a conduct analysis. Here, the role of the legal standard is to define what conduct is rightful and wrongful.

When we take into account this distinction, it appears that, in some jurisdictions, private nuisance is a strict liability tort. As Professor Keating argues, this is most clearly demonstrated by *Boomer v Atlantic Cement*.¹⁸⁶ In that case, neighbors of a nearby cement factory complained that the smoke, dirt, and vibrations emanating from the factory constituted a nuisance. The court imposed damages but refused to impose an injunction. The factory employed three hundred people. Forcing it to close would have resulted in great unemployment. According to Professor Keating, the plaintiff in this case had a right to reasonable use and enjoyment of the land.¹⁸⁷ This right had been infringed and therefore liability in the form of damages was imposed. However, the defendant's conduct was not wrongful. The defendant's conduct was socially beneficial, and thus reasonable. As a result, the court refused to enjoin the production. Nevertheless, the court fundamentally

185. See *supra* note 54.

186. 26 N.Y.2d 219 (1970).

187. Keating, *Nuisance*, *supra* note 18, at 42–50.

applied a strict liability rule. The plaintiff had been affected in a legally relevant way (the right to reasonable use and enjoyment had been infringed). Therefore, liability followed, regardless of the fact that the conduct was not wrongful, and the defendant was required to pay damages for infringing a right.

Returning to copyright, the question posed is whether the fairness standard facilitates an impact analysis or a conduct analysis. In asking whether the copying was fair, does the court primarily ask: was the defendant impacted in a legally cognizable fashion? Or, is the court's question primarily: was the defendant's copying good or bad for society? If it is the former, then the fairness standard is relevant only to the issue of defining the scope of the right. If it is the latter, the role of the fairness standard is to separate rightful from wrongful conduct.

The answer is that the fairness standard is primarily a conduct analysis. As the balancing of interests theory highlights, the ultimate and final arbiter in fair use is the question of whether the defendant's copying is socially beneficial. The root of the fair use analysis is an examination of whether the defendant's conduct will produce greater access-benefits or incentive-costs. Conduct that produces negative consequences are labeled unfair; conduct that produces positive consequences are labeled fair. Thus, the fairness analysis seems primarily concerned, not with whether a right was infringed, but with whether the conduct was good for society and thus not faulty. The fairness standard therefore does not define the scope of the right, but distinguishes between rightful and wrongful conduct. As a result, copyright infringement is fault-based. Unlike nuisance, which imposes liability whenever the defendant's actions impact the plaintiff in a legally cognizable fashion, liability is only imposed in copyright if the copying is unfair (i.e., if this conduct is bad for society).

However, once again, this analysis is complicated by the presence of the harm concept in the fair use doctrine. While the question of fairness is primarily a question of conduct, the fairness analysis does contain an impact analysis within it, as demonstrated by cases such as *Campbell*. These cases are ones in which no liability follows because the court determines that the plaintiff was not impacted in the legally relevant fashion.

The fact that harm slips into the fault analysis explains the statutory wording in the Copyright Act. The Copyright Act currently says that fair copying is not an infringement of the right. This is half-correct. As harm finds itself in the fair use inquiry, some fair use cases, such as *Campbell*, represent instances where the copying causes no harm, and therefore no right has been infringed. The drafters' only mistake was simply a failure to write the statutory language in a sufficiently nuanced manner that would express

the dual character of fair use. While some instances of copying are fair because they involve no harm and thus no rights-infringement, other instances of copying are fair because, although they do involve harm and accordingly a rights-infringement, that harm is nonetheless justifiable because the copying is good for society.

2. *Fair Use is an Affirmative Defense*

The second critique arises out of the fair use doctrine's status as an affirmative defense. As a result, even if fair use is a fault inquiry, the plaintiff is not required to prove fault as part of the prima facie case; instead, the defendant must prove the absence of fault. On this basis, some may argue that the liability rule is in fact strict. Because the issue of fault only becomes relevant to the liability decision if the defendant pleads fair use, defendants who do not plead fair use will be held liable on the basis of copying and substantial similarity alone. It would appear therefore that fault is not a necessary condition for liability and frequently defendants are held liable even when no one has introduced any evidence about fault.

To respond to this claim, this Subsection shall first demonstrate that the position of fair use as an affirmative defense in fact strengthens the claim that copyright is a fault-based tort. Second, it shall then proceed to try to accurately characterize the procedural role of fault in copyright infringement.

a) Fair Use as an Affirmative Defense

Recall that the types of affirmative defense differ depending on whether the tort is judged by a strict liability rule or a fault liability rule. In strict liability cases, only defenses of plaintiff fault exculpate the defendant. Conversely, justification and excuse are admissible defenses in cases of fault liability. The question copyright scholars must ask, therefore, is what type of defense is fair use?¹⁸⁸

To answer this question, we may begin by demonstrating what a fair use claim clearly is not. First, it is not a claim that the plaintiff was at fault for the outcome. Defenses like contributory negligence deny the plaintiff relief because he was at fault for the accident. In copyright, a similar defense would be to demonstrate how some conduct of the copyright holder was faulty and resulted in the copied work. But clearly the question of fair use centers not on the actions of the copyright holder, but on the actions of the copyist.

Nor is fair use an excuse. Recall that excuses focus on the subjective characteristics of the defendant. They argue that, although the conduct was

188. On other limitations and exceptions, see *supra* note 141.

wrongful, and there was fault in the action, the defendant did not act with any bad will, and therefore there is no fault in the actor. Clearly fair use does not fall into this category, as the impact of a defendant's mental state is not relevant to the fair use determination.

If fair use is an affirmative defense, then it is doubtlessly a justification. Like other justifications, its purpose is to demonstrate that the defendant's conduct was not wrongful. Rather fair use is rightful conduct. Fair copying produces benefits of access that outweigh any negative consequences it may have. Thus, by introducing a fair use claim, the defendant argues that there was no fault in the action, and instead demonstrates that this was actually good conduct. Therefore, if fair use is an affirmative defense, this simply provides another avenue for proving the same thing: that liability ultimately depends upon the defendant's fault.

b) The Procedural Role of Fault in Copyright Infringement

Nevertheless, there is something unusual about copyright's formal structure. Normally in a fault-based tort, the plaintiff must introduce at least some evidence establishing that the defendant was at fault for the outcome. The burden then shifts to the defendant to defend himself through excuse or justification. In copyright infringement, the copyright holder need not provide any evidence of the copyst's fault. He must only demonstrate conduct and outcome. Thereafter, the defendant can absolve responsibility by claiming the absence of fault under the paradigm of fair use. Although copyright is a fault-based tort, it is unusual because there is no burden on the copyright holder to prove any unfairness, only a burden on the copyst to prove fairness.

It would seem, therefore, that in a copyright action, the court apparently presumes the existence of fault. Copying is presumptively unfair until the copyst can be shown otherwise. This itself is not necessarily unusual. There are other situations where the court will presume the existence of fault until the defendant can rebut that presumption. This happens most classically under the *res ipsa loquitur* doctrine in negligence.¹⁸⁹ This doctrine, which in English means "the thing speaks for itself," is applied in cases where in all probability the accident could not have occurred without fault on part of the defendant. It is used commonly today in surgical malpractice cases.¹⁹⁰ Where

189. RESTATEMENT (SECOND) OF TORTS, § 328; *Byrne v. Boadle*, (1863) 159 Eng. Rep. 299 (Exch.); 2 Hurl. & Colt. 722.

190. *See, e.g., ZeBarth v. Swedish Hosp. Med. Ctr.*, 81 Wash. 2d 12, 18 (1972); *Douglas v. Bussabarger*, 73 Wash. 2d 476 (1968); *Leach v. Ellensburg Hosp. Ass'n, Inc.*, 65 Wash. 2d 925 (1965).

a foreign body is found in a patient following a surgery, there is a very high likelihood that the surgeons were negligent, and accordingly the court will simply presume the fault unless the defendant can rebut that presumption. A similar presumption is found in defamation by libel.¹⁹¹ Proving a case of defamation at common law traditionally required that a defendant publish some defamatory material concerning the plaintiff. More recently courts have also required that the statement be false, that there be some degree of fault on the defendant's part, and that it caused actual damages.¹⁹² However, in cases where the defamation is in print, the courts have presumed fault and actual damages on the part of the defendant. In such cases, it is presumed that the defendant acted maliciously.¹⁹³ It is then up to the defendant to rebut that presumption at the defense stage.

However, characterizing fair use as an attempt to rebut a presumption of fault is problematic. Presumptions are exceptions to the normal rule that the plaintiff must prove conduct, outcome, and fault. In order for the court to apply this exception, the plaintiff must typically demonstrate some additional, supplementary factual condition. For example, in negligence, the plaintiff must typically prove conduct, outcome, and fault. However, in a subset of cases, the plaintiff can argue that he suffered a type of harm that does not normally occur without negligence, then he will gain the benefit of the *res ipsa loquitur* doctrine. Likewise, in defamation, usually the plaintiff must prove fault. However, if the plaintiff can make the additional showing that the defamatory material was published in print, then the court will presume fault.¹⁹⁴ In both cases, the presumption only operates after the plaintiff has demonstrated a reason why it should operate.

Copyright infringement does not easily fall within this mold. Once the copyright holder proves copying and substantial similarity, the burden of proving fairness is always placed on the copyist. If copyright adopted a rebuttable presumption, then typically the copyright holder would be required to prove copying, substantial similarity, and unfairness, unless he could prove some supplementary condition which would justify making the copyist prove fairness in a subset of cases.

In either case, it appears that copyright infringement is somewhat anomalous within the broader field of tort law. If fair use is an affirmative defense, then copyright infringement is unusual in that the plaintiff need not introduce evidence to prove the existence of fault before the defendant must

191. DOBBS, *supra* note 2, § 401.

192. *Id.*

193. *Id.*

194. *Id.* (on presumption of fault in libel).

offer a defense. Alternatively, if fair use serves to rebut a presumption of fault, copyright infringement takes the unusual position of placing the burden of proving fault on the plaintiff until he introduces evidence showing why that burden should be shifted.

3. *A Caveat: The Market Failure Approach to Fair Use*

Finally, before closing this Part, we must introduce one caveat into the analysis. So far, this Article has sided with Professor Lunney's interpretation that fair use is ultimately a consequentialist balancing test. If so, the thesis that copyright infringement is a fault-based tort is wholly plausible. A plaintiff's right is infringed when copying leads to a substantially similar work that displaces demand (i.e., when the proscribed conduct causes a harmful outcome). Such rights-infringement will, nevertheless, not result in liability in instances of fair use because, in such cases, the defendant's conduct was actually a good thing for society, and accordingly not wrongful: it is fair.

However, if the older market failure approach is correct, then the thesis is more dubious and copyright is probably a strict liability tort. Crucially, in the market failure theory, whether the copying was socially desirable is not determinative as to whether fair use should apply. Instead, if a defendant engages in socially desirable copying, the expectation is that he should still pay the copyright holder for the ability to do so (either in a license fee or in damages). This normal rule is displaced only in the case where a market failure occurs that would prevent the socially desirable copying. This would not only be a strict liability rule, but it would be the perfect example of a strict liability rule.

As noted above, strict liability does not take rightfulness or wrongfulness into account when imposing liability; all that matters is whether a right has been infringed. Thus, we see classic cases where the court accepts that the conduct was rightful, or not wrongful, but, because liability is strict, it imposes liability anyways. For example in the infamous trespass case of *Vincent v Lake Erie*, the court accepted that the defendant's action of tying his ship to the plaintiff's dock was not wrongful (it was instead socially beneficial), but held that nevertheless the resulting harm to the defendant's property amounted to an infringement of his right and therefore that the plaintiff ought to be compensated.¹⁹⁵

If the market failure approach to fair use is descriptively accurate, then the same situation occurs in copyright. Even in the case where the defendant's copying is in fact socially desirable because the benefits of access

195. 109 Minn. 456 (1910).

outweigh harm to incentives, the defendant will nevertheless be required to pay the plaintiff. Thus, in this interpretation of the liability rule, even rightful conduct that is good for society and not wrongful results in liability. This is the canonical case of strict liability. Nevertheless, this Article has based its opinion on the newer balancing approach to fair use, which has critiqued the market failure approach.

IV. RESTRUCTURING FAIR USE

So far this Article has focused on an analytic exercise: our goal has been to determine whether copyright infringement adopts a strict liability or a fault liability rule. The thesis presented is that copyright adopts a fault rule, similar to a negligence rule. But why does this characterization matter? What is the real world relevance of labeling copyright strict or fault-based?

The answer to this question is twofold. Firstly, a number of scholars have argued that copyright's supposed adoption of a strict liability rule is normatively undesirable, and therefore they recommend that copyright be altered to a fault-based regime. In particular, frequent claims are made that copyright infringement ought be restructured as an intentional tort. However, as this Article has argued, it appears that copyright infringement is in fact a fault-based tort in the same way that negligence is. As a result, copyright's liability regime may not be quite as bad as previous scholars have suggested. As this Part demonstrates, the type of fault required before liability is imposed in copyright is currently optimal, given copyright's underlying normative structure.

Secondly, as Part III demonstrated, the formal structure of copyright infringement is a mess. The most obvious example of this mess is the way harm and fault collapse into one another in the fair use analysis. The second Section of this Part elaborates on the two problems this collapsing causes. Firstly, it results in courts trying to reach findings of no-liability by appealing to the concept of "no harm," when they really mean that there is no fault. Secondly, it results in the burden of proof being poorly assigned. It is therefore recommended that courts try to distinguish these concepts. This Part concludes by demonstrating how the fair use analysis could be restructured to accomplish this end.

A. THE NORMATIVE DEFENSIBILITY OF COPYRIGHT INFRINGEMENT'S LIABILITY RULE

This Section first summarizes the criticisms that are often presented against the supposed strict liability rule in copyright before showing that these criticisms are overstated given that copyright infringement is already based upon fault. It then goes on to evaluate whether copyright infringement

would be normatively improved if it were to be reformed as an intentional tort.

1. *The Normative Critique*

For decades, academicians have offered criticisms of the supposed strict liability rule. In answer to these criticisms, scholars typically recommend that copyright be altered to base liability upon some element of fault. We briefly summarize these critiques here.

a) Inconsistency

First, it is argued that reliance on strict liability is anomalous within the greater field of tort law.¹⁹⁶ The standard historical account of the common law states that the early law was based on strict liability, but over time tort has gradually replaced strict liability rules with fault liability rules. In particular, many accounts point to the mid-nineteenth century as the period when the common law moved from a regime based primarily on strict liability to becoming a regime primarily based on negligence. It is often said that the case of *Brown v. Kendall* is a key point within this evolution.¹⁹⁷ In an oft-quoted passage from this case, Chief Justice Marshall of the Massachusetts Supreme Court decided that in order for one to be liable for harm accidentally caused to another, it must be shown that the defendant failed to take “the kind and degree of care, which prudent and cautious men would use, such as is required by the exigency of the case.”¹⁹⁸ Clearly strict liability did not die off after this case,¹⁹⁹ but the change in attitude that it represented left us with a modern regime in which the bulk of tort liability is assessed through the use of negligence rules.²⁰⁰ Why then is copyright any different? In a world where the common law has generally moved away from

196. See, e.g., Hetcher, *supra* note 15, at 1283–99 (describing the historical evolution from strict liability to fault liability in tort generally), 1290 (“My present concern is not whether these are convincing arguments from a top-down normative perspective. For present purposes, what matters is that the fault standard won out in the case law. This doctrinal development is of interest because it raises the obvious question: if a heightened moral sensitivity toward fairness and social welfare dictated a move to the fault standard in tort generally, then why not in copyright as well?”).

197. 60 Mass. 292 (1850).

198. *Id.*

199. Briefly, in the nineteen seventies, strict liability made a reappearance in the theory of enterprise liability. See WHITE, TORT LAW IN AMERICA, *supra* note 1, at 168–72.

200. COLEMAN, *supra* note 4, at 218 (“The bulk of fault liability involves negligence.”); Richard A. Posner, *A Theory of Negligence*, 1 J. LEGAL STUD. 29, 29 (“negligence cases[] constitute the largest item of business on the civil side of the nation’s trial courts”).

strict liability towards fault liability, what justifies copyright's decision to maintain a strict liability standard?²⁰¹

b) Inefficiency

Second, there is a concern that strict liability in this context may deter some cases of good, beneficial copying.²⁰² As previously discussed, some copying is simply good conduct because it creates benefits of greater access. Ideally, a social planner would wish to encourage this behavior. However, copyright law is notoriously complicated. Not only does the 1976 Copyright Act contain over thirteen hundred sections (a length that makes it comparable to the tax code), but the conceptual difficulty of dealing with intangible goods has led to copyright (together with patents) being called “the metaphysics of the law.”²⁰³ For the ordinary citizen, it is often very difficult to assess whether they have copied enough protected expression to infringe the copyright holder's exclusive right. In this context, we may see over-deterrence. Some may forgo copying that is beneficial because they cannot accurately assess whether the copying is lawful. This is exacerbated by the risk-aversion that many people demonstrate. Previously some scholars have suggested that this problem would be alleviated if liability turned on some question of fault.²⁰⁴ If the defendant could argue that he did not copy intentionally, recklessly, or negligently, then he may feel more confident in engaging in beneficial copying.

201. Professor Reese argues that U.S. copyright law drifted into this state of strict liability by accident. See R. Anthony Reese, *Innocent Infringement in U.S. Copyright Law: A History*, 30 COLUM. J.L. & ARTS. 133 (2007) (arguing that traditional protections offered to innocent infringers gradually were stripped away during the twentieth century as a byproduct of other changes occurring in U.S. copyright law).

202. See, e.g., Ciolino & Donelon, *supra* note 15, at 413 (“Strict liability overdeters lawful and beneficial uses of copyrighted works.”); Reese, *supra* note 201, at 183 (“Because copyright law seeks to encourage such noninfringing copying, the possibility of holding innocent infringers liable should be worrisome if it deters potential users from using copyrighted material in ways that might ultimately be found noninfringing”); Eva Subotnik, *Intent in Fair Use*, 18 LEWIS & CLARK L. REV. 935, 963–64 (2014) (“Users (even those with legal counsel) often find themselves unable to predict with confidence whether a use would be deemed fair . . . and risk aversion will lead some to abandon projects rather than come close to the boundary line between fair use and infringement.”).

203. *Folsom v. Marsh*, 9 F. Cas. 342, at 344 (C.C.D. Mass. 1841).

204. See, e.g., Ciolino & Donelon, *supra* note 15; Subotnik, *supra* note 202, at 976 (proposing that incorporating intent into fair use determinations will alleviate problems of over-deterrence and “depriving the public of socially beneficial uses.”)

c) Immorality

Finally, there is an argument that strict liability in copyright is simply unfair. In the past, some deontological scholars have tried to demonstrate the immorality associated with strict liability rules. Professor Jules Coleman has argued that the “substitution of fault for causation marked an abandonment of the immoral standard of strict liability under Trespass (which, after all, imposed liability without regard to fault) in favor of a moral foundation for tort law based on the fault principle.”²⁰⁵ Moreover, Professor Ernest Weinrib argues that strict liability creates an unjust inequality between the plaintiff and defendant. In this view, strict liability reflects “extreme solicitude for plaintiffs’ rights” with little weight given to the defendant’s equal interest in living an autonomous life.²⁰⁶ In the copyright context, Professor Dane Ciolino and Erin Donelon have argued that copyright’s strict liability regime “conflicts with traditional deontological notions of personal autonomy.”²⁰⁷ By requiring copyists to pay damages for actions that they did not intentionally cause, copyright forces the individual to bear the responsibility for consequences that they have not willfully brought about. Instead they argue copyright should only hold individuals liable if they copy intentionally.

2. *Answering the Normative Critique*

The fact that copyright is not based on strict liability forces us to reconsider these criticisms. While this Article does not suggest that the rules governing copyright infringement are currently without flaw, the fact that they already require some element of fault reduces the impact of these normative arguments.

a) Inconsistency

Arguably the most misplaced of critiques is that copyright’s reliance on strict liability is inconsistent with the rest of tort doctrine. Most torts require the defendant to act with fault before liability will be imposed. But even more salient is the fact that liability for most torts requires only negligent conduct. That is, in most cases, the fault is not based on the defendant’s state of mind, but on whether he failed to comply with a standard of conduct. With this in mind, copyright’s liability rule, which also requires the defendant

205. Coleman, *supra* note 4, at 374.

206. WEINRIB, *supra* note 4, at 179.

207. Ciolino & Donelon, *supra* note 15, at 419. In addition, other authors who have talked about the “harshness” of strict liability seem to suggest there is some unfairness involved in the current liability regime. See, e.g., Mullally, *supra* note 129; Depoorter & Walker, *supra* note 129.

to fail to comply with a standard of conduct before imposing liability, seems not anomalous, but largely consistent with the broader field of tort doctrine.

b) Inefficiency

Perhaps most important, though, is the demonstration that copyright's liability rule is broadly efficient. The over-deterrence argument suggests that currently copyright produces incentives to act in inefficient ways (i.e., by forgoing economically beneficial copying). The analysis provided here, however, suggests a different story. If copyright adopted a strict liability rule, much efficient copying with great benefits in terms of access would be forgone. However, the fault inquiry that lies at the heart of fair use exculpates defendants when their copying is beneficial for society. The law is organized in such a way that economically beneficial copying does not result in liability.

This is not to say that over-deterrence does not happen. It is still highly possible that, due to the complexity of copyright, users of copyrighted works will be unable to accurately determine whether their copying is lawful or not and, as a result, may shy away from copying that would benefit society. However, what the analysis does reveal is that this is not a problem with the liability rule *per se*. If individuals act in conformity with the liability rule (copying when doing so is fair, refraining from doing so when it is not), then efficiency will be reached. People behave inefficiently not because the liability rule in place is inefficient, but because they do not fully understand what the liability rule requires of them. The complexity of copyright makes it difficult to determine whether they are acting in conformity with the standard the law establishes. This encourages people to shy away from uses that, while lawful, may be approaching the border between infringement and fair use.

Given that this is the case, the appropriate response is not to change an already efficient liability rule, but to better educate people of their duties established by the law. Informing people more clearly on what is a copyright infringement and what is a fair use will lead people to acting in conformity with the efficient liability rule that copyright infringement already adopts.²⁰⁸ To that end, the promulgation of fair use guidelines is particularly important.²⁰⁹ By establishing and distributing such guidelines, we can instill

208. Assuming that individuals try to follow formal law, which is perhaps a debatable assessment of behavior in this area. *See generally* TINY MURRY ET AL., *PUTTING IP IN ITS PLACE* (2014).

209. *See generally* Gregory Klingsporn, *The Conference on Fair Use (CONFU) and the Future of Fair Use Guidelines*, 23 COLUM.-VLA J.L. & ARTS 101 (1999); Robert Thornburg, *The Impact of Copyright Law on Distance Education Programs: How Fair Use and the CONFU Guidelines May Shape the Future of Academia*, 27 S. ILL. U. L.J. 321 (2003). The work of Professor Peter Jazi and Professor Patricia Aufderheide on improving fair use awareness in the documentary

some confidence in those who wish to copy for lawful and beneficial purposes.

c) Immorality

The fact that copyright is based on fault also demonstrates that our test for copyright infringement is not as immoral as perhaps once thought. Professor Weinrib's argument that strict liability offers "extreme solicitude" for plaintiff's rights without equally taking into account the legitimate interests of defendants is undoubtedly true in many instances, but it is not applicable in the copyright context. As demonstrated, the fair use doctrine applies in a multitude of highly diverse factual situations to protect the interests of the copyist. Whether the law upholds the interests of the right holder or the copyist depends not on some unjust favoritism, but on an objective determination about how to bring about the greatest social benefit.

Equally, Ciolino and Donelon's argument that strict liability in copyright fails to take seriously the notion of personal autonomy apparently forgets that the law often holds people liable for actions they did not intend. Defendants in negligence cases are frequently held liable, although they have not willfully brought about the harm they cause. If holding a defendant liable for unintentional copying is immoral because it fails to respect people as autonomous beings, then it is at the very least no more immoral than the large swathes of tort law that hold defendants liable for their unintentional but nevertheless negligent actions.

3. *Reforming Copyright Infringement as an Intentional Tort?*

Although copyright's liability regime may not be as inconsistent, inefficient, or immoral as previously has been suggested, that does not mean it is without flaw. While the liability rule may not be as bad as once supposed, there could still potentially be room for improvement.²¹⁰ Those who have

filmmaking community is also worthy of special praise. See Peter Jaszi & Patricia Aufderheide, *Untold Stories: Collaborative Research on Documentary Filmmakers' Free Speech and Fair Use*, 46 CINEMA J. 133 (2007); see AUFDERHEIDE ET AL., STATEMENT OF BEST PRACTICES IN FAIR USE OF COLLECTIONS CONTAINING ORPHAN WORKS FOR LIBRARIES, ARCHIVES, AND OTHER MEMORY INSTITUTIONS (2014), available at <http://www.cmsimpact.org/sites/default/files/documents/orphanworks-dec14.pdf> (last visited Feb. 27, 2015).

210. One flaw with the current liability rule comes in its application. In applying the fairness rule, courts assume that the a copyist can tell with certainty whether copying will harm incentives. As a result, when applying the fairness rule, courts compare simply the access-benefits with the incentive-costs. However, the assumption that the copyist can tell the incentive effects ex ante is unrealistic in many cases. Typically, the would-be copyist can only tell that there is a certain probability that the use will create some harm. Ideally in these cases, the court's assessment of the fairness of copying should not depend on an ex post

researched this topic in the past have usually suggested that copyright infringement become an intentional tort. Hence, Professor Ciolino and Donelon argue that the copyist's lack of intention should be a complete defense to copyright infringement.²¹¹ Professor Jacqueline Lipton acknowledges this as one potential route, but also suggests that the plaintiff must prove intent as part of the prima facie case.²¹²

This raises the question, what category of fault should copyright liability be based upon? It currently is based upon the failure to comply with a standard, but would the situation become normatively superior if liability were to be based upon the defendant's mental state? In particular, ought copyright infringement be reformed into an intentional tort? In response to these suggestions, this Article takes the view that the status quo ought to be maintained.

a) What Does Intent Mean in Copyright?

As an initial matter, we must firstly be careful to clarify what the concept of intent would mean in this context. Professor Eva Subotnik has wisely pointed out that “[t]hose who invoke user ‘intent’ in these contexts are often not precise in what they mean by that term and related concepts, such as ‘good faith.’”²¹³ Professor Subotnik then proceeds to isolate three different types of intent that could potentially be relevant to the liability decision. Those are: the “intent to communicate new meaning, intent to comply with the law of fair use, and intent to be a good citizen.”²¹⁴ However, while this Article applauds Professor Subotnik's serious engagement with a difficult analytic question that has gone under-theorized, it does not agree that these are the only types of intent that are relevant in copyright law, and in fact believes that her article has overlooked the most relevant definition of intent for copyright purposes.

As noted earlier, the intent concept in tort law, especially when used in reference to fault liability, connects the elements of conduct and outcome. In this manner, it is distinguishable from similar concepts such as volitional or

comparison of access-benefits and incentive-costs. Instead, courts should engage in an ex ante comparison of the access-benefits with the *expected* incentive-costs (where expected incentive-costs are defined as the potential harm to incentives multiplied by the probability of their occurrence). For a more detailed treatment of this issue, see Oren Bracha & Patrick R. Goold, *Copyright Accidents or: Should Copyright Have A Negligence Standard?* (unpublished draft on file with authors).

211. See Ciolino & Donelon, *supra* note 15, at 410, 420.

212. Lipton, *supra* note 15, at 804.

213. Subotnik, *supra* note 202, at 947.

214. *Id.*

deliberate conduct. “Intentional torts” are fault-based because the defendant has engaged in the prescribed conduct with the purpose of causing the legally cognizable form of harm.

The conduct in copyright is copying, while the harmful outcome is market harm flowing from the existence of the substantially similar work. Therefore, if copyright infringement were to become a fault-based tort based upon intention, liability would require that the defendant engage in copying with the intention of causing market harm. The intentional tort of copyright infringement would impose liability upon a defendant only when it is proved that she actually tried to divert the copyright holder’s sales or reasonable license fees. The question is whether such a modification would be normatively desirable.

b) Ought Copyright Infringement Be an Intentional Tort?

This Article takes the position that copyright ought not become an intentional tort. While this may potentially improve the morality of the copyright system, doing so would interfere with the efficient liability regime already in place.

We have previously seen that the liability rule in copyright is broadly efficient. If the balancing interpretation of fair use is correct, then the law creates incentives for users to engage in copying when doing so produces greater benefits in terms of access than cost in lost incentives, and to refrain from copying in cases where such copying would cause harm to incentives that is not offset by benefits of access. As a result, we have a system that (at least theoretically) serves overall social welfare.²¹⁵

Reforming copyright infringement into an intentional tort would jeopardize the efficiency that the current system creates. Crucially, whether a defendant’s copying is good or bad from a welfare perspective is totally unaffected by the defendant’s mental state. When assessing whether copying is good or bad from a social perspective, all we need to assess is whether the copying harms incentives greater than it benefits access. The upshot is that if we exculpate defendants who create unfair copies on the basis that they did not intend to produce the legally cognizable harm, then we permit individuals to create copies that will harm future incentives without any greater

215. One small caveat ought to be highlighted. So far, this Article has discussed the liability rule in copyright “under the simplifying assumption that there is a certainty rather than a probability of harm” flowing from the defendant’s conduct. If the link between the conduct and the harmful outcome is certain, then the liability rule is efficient. It is not clear whether this would be the case if this assumption were to be relaxed. *See* LANDES & POSNER, *supra* note 55, at 29, 54.

offsetting benefit of public access. This is strongly antithetical to the normative goal of promoting social welfare.²¹⁶

On the other hand, reforming copyright infringement into an intentional tort would arguably improve the morality of the system. As noted, the standard of conduct type of fault is frequently defined in consequentialist terms. Economic scholars have provided very few compelling reasons as to why intentional conduct is tortious and not simply dealt with by the criminal law.²¹⁷ Deontological scholars, by contrast, have the reverse problem. Arguing from Kantian ethics, it is, to some degree, intuitively demonstrable that intentionally causing harm is wrongful. But these scholars have found it much harder to explain negligence in such terms.²¹⁸

If copyright scholars wish to see copyright adopt an intentionality requirement, the best possible argument to make would rest on the deontological position that whether an action is right or wrong depends on the will of the actor, not its consequences. Thus, the appropriate normative basis for holding the defendant's copying liable would be found in the blameworthy state of mind this represents. Nevertheless, such an argument would seem inappropriate in the Anglo-American context, given the frequent claims that the consequentialist and utilitarian goal of maximizing social welfare informs the entire system.²¹⁹ In particular, if the court begins to use

216. This concern has already been highlighted. See Michael Madison, *A Pattern-Oriented Approach to Fair Use*, 45 WM. & MARY. L. REV. 1525, 1555–56 (2004) (“[W]hy the ‘good faith’ of the infringer should matter here is unclear. To the extent that copyright policy is informed by a utilitarian calculus maximizing social welfare in terms of ‘creativity’ and ‘creative’ works of authorship,” the question is not “whether the defendant believed that he or she was acting legitimately,” but is instead “whether the outcome of the defendant’s efforts was more socially valuable than the outcome produced by allowing the copyright holder to enjoin the use or obtain payment.”). Although, given the difficulties courts face in calculating the incentive-cost and access-benefit variables, room may still exist to employ the defendant’s intent as a rough heuristic device for ascertaining the social welfare impact of the copying.

217. See *supra* note 82.

218. See generally Heidi M. Hurd, *Finding No Fault with Negligence*, in THE PHILOSOPHICAL FOUNDATIONS OF THE LAW OF TORTS 387 (John Oberdiek, ed., 2014); Alan Calnan, *The Fault(s) in Negligence Law*, 25 QUINNIPIAC L. REV. 695 (2007).

219. See, e.g., *Twentieth Century Music Corp. v. Aiken*, 422 U.S. 151, 156 (1975) (“The immediate effect of our copyright law is to secure a fair return for an ‘author’s’ creative labor. But the ultimate aim is, by this incentive, to stimulate artistic creativity for the general public good.”); *Mazer v. Stein*, 347 U.S. 201, 219 (1954) (“The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in ‘Science and useful Arts.’”); *United States v. Paramount Pictures, Inc.*, 334 U.S. 131, 158 (1948) (“The copyright law, like the patent statutes, makes reward to the owner a secondary consideration. . . . It is said that

deontological premises to define what counts as an actionable wrong in this area, coherency would require using deontological premises to justify the initial grant of right.²²⁰ While such a prospect is hardly beyond the bounds of possibility, it would cut against the frequent statements made by academics that copyright is not about natural right, but instead about social welfare.²²¹

B. THE FORMAL STRUCTURE OF COPYRIGHT INFRINGEMENT

While the liability regime need not require any modification given copyright's underlying normative goals, there is a strong case for altering the formal structure of copyright infringement analysis. As this Article has shown, the role of fault in copyright infringement is often obscured by the collapsing of harm and fault inquiries into the fair use analysis. As this Section elaborates upon, this results in "no fault" cases being shoe-horned into "no harm" language, and poorly assigned burdens of proof. After which, this Section demonstrates how restructuring the fair use analysis could address both of these problems.

1. *Collapsing Harm and Fault*

Theoretically, the fact that fair use contains two separate conceptual inquiries is not itself problematic. As long as judges are aware that asking whether harm occurred and whether the harm was justifiable are two distinct and equally important questions, then it does not particularly matter whether both are discussed under the doctrinal label of fair use or disaggregated into their own separate doctrines.

However, in practice, the dual nature of fair use inquiry does lead to a problem. Professor David Nimmer previously highlighted the problem of "stampeding" in fair use cases.²²² By this he meant that judges make their

reward to the author or artist serves to induce release to the public of the products of his creative genius.").

220. This is certainly not to say that scholars in the Anglo-American tradition are completely blind to deontological foundations for intellectual property rights. See e.g., ROBERT MERGES, JUSTIFYING INTELLECTUAL PROPERTY (2011) (justifying intellectual property rights from Lockean, Kantian, and Rawlsian perspectives); Justin Hughes, *The Philosophy of Intellectual Property*, 77 GEO. L.J. 287 (1988).

221. See, e.g., Stephen Breyer, *The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 HARV. L. REV. 281, 291 (1970) ("In sum, none of the noneconomic goals served by copyright law seems an adequate justification for a copyright system.").

222. David Nimmer, *"Fairest of them All" and Other Fairy Tales of Fair Use*, 66 LAW & CONTEMP. PROBS. 263, 281 (2003) ("Courts tend first to make a judgment that the ultimate disposition is fair use or unfair use, and then align the four factors to fit that result as best they can. At base, therefore, the four factors fail to drive the analysis, but rather serve as

conclusion as to whether the use is fair and then use the four factors to justify their conclusion. The fair use factors do not drive the analysis, but instead are used to support the conclusion. This is particularly noticeable when it comes to the fourth factor on fair use. Because the presence of market harm has been considered the most important factor in the fair use analysis, judges who wish to find copying to be fair had better work particularly hard to show that this vital factor does not go against their overall conclusion.

But as noted earlier, if the balancing test for fair use is accurate, then the touchstone of fair use analysis is not whether the defendant caused harm, but whether that harm was justifiable. There are cases where the defendant causes the legally cognizable market harm, but this should not attract liability because that copying was still a good thing for society. However, in such cases, if a judge is subject to the stampeding effect, then it is likely he will try to align the fair use factors in a way that will suit his conclusion. Accordingly, there will be pressure for him to conclude that the fourth factor on market harm supports his finding of no liability, or at the very least does not cut against it. Therefore, in some cases where there is harm but that harm is justifiable, it seems plausible that judges will try to justify their finding of no liability by appealing to the concept of “no harm,” rather than the concept of “no fault.”

Perhaps the most salient example of this occurring comes in the Google Images controversy. As discussed earlier, there was a strong claim that Google’s Image Search caused the plaintiff some cognizable legal harm because it arguably displaced sales in the cellphone market. Nevertheless, despite the arguable case of harm, the Ninth Circuit said that this harm “remain[ed] hypothetical” and thus concluded that the fourth factor favored neither side.²²³ In doing so, the court appealed to the lack of empirical proof of harm. Yet such reasoning is questionable. Not only does the fourth factor only require harm to a “potential” market, but there was a theoretical reason to believe such harm was occurring.

This author interprets this case as one of stampeding. The court recognized that there was significant public benefit to Google’s use and that, all things considered, the benefit in terms of access outweighed the harm caused to Perfect 10 and the lost future incentives this may cause. Nevertheless, because of the stampeding effect, the court believed that they could not justify a conclusion of no liability without at least demonstrating

convenient pegs on which to hang antecedent conclusions.”). *But see* Beebe, *supra* note 107, at 588–94.

223. *Perfect 10*, 508 F.3d at 1168.

that the market harm factor did not cut against their overall verdict. There was an arguable case of market harm, but rather than acknowledge the harm and state clearly that this was a case in which the benefits outweighed the harm, the court tried to reason away the harm by appealing to the lack of empirical evidence.

On their own, isolated instances of such stampeding are unlikely to be too problematic. However, if this sort of reasoning is commonplace amongst the judiciary, then a serious consequence emerges. Routinely fitting “no fault” cases into the language of “no harm” creates a dangerous precedent. As the language of “no harm” increasingly dominates the body of fair use cases, the more pressure future litigants will experience to likewise fit their claim into such “no harm” language. This prejudices those defendants who engage in copying that does cause harm, but which is nevertheless socially desirable. The lack of fault in these cases should be enough to ground a conclusion of no liability. However, because there is a pressure to fit claims into the language of “no harm,” defendants will need to formulate rather tenuous arguments as to why their use is not harmful. Indeed, we may already be in this position. In *Perfect 10*, Google tried to fit its “no fault” claim into the language of “no harm,” but as there was a strong reason to believe harm existed in this case, the trial court dismissed its legitimate case. Only through rather questionable reasoning did the Ninth Circuit salvage Google’s Image Search.²²⁴

224. Arguably stampeding can be seen in other legitimate “no fault” cases. In *Bill Graham Archives*, the Second Circuit found no market harm. The court reasoned that the defendant’s use was transformative, and that the plaintiff could not prevent others from entering transformative markets. *Bill Graham Archives v. Dorling Kindersley, Ltd.*, 448 F.3d 605, 614–15 (2d Cir. 2006). Again, this seems questionable. Under *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994), the test is whether the defendant’s use harms licensing market that are “traditional, reasonable, or likely to be developed.” In the *Bill Graham* case, the plaintiff was already in the market for licensing reduced sized versions of the posters. The case for market harm was at least plausible. Likewise, in *Authors Guild v. Google*, the court held that snippet views did not cause market harm. Firstly, the court reasoned that it was unlikely that “someone would take the time and energy to input countless searches to try and get enough snippets to comprise an entire book.” *Authors Guild, Inc. v. Google Inc.*, 954 F. Supp. 2d 282, 293 (S.D.N.Y. 2013). However, such reasoning forgets that copyright infringement can occur when someone reproduces a qualitatively important, yet quantitatively small, section. For example, if *Hamlet* were still under copyright, and someone were to reproduce part of the famous “to be or not to be” soliloquy, this would likely amount to an infringement despite the relatively small amount of copying. Likewise, Google’s snippet views may well allow people access to qualitatively important parts of text, resulting in market substitution. Secondly, the court reasoned that Google Books allows authors to become noticed, and thus may increase sales as demand increases for the work. *Id.* However, there are a number of problems with such reasoning. Most fundamentally, as a matter of private law, it is not usually the case that the defendant’s

2. *The Burden of Proof*

Finally, as the fair use doctrine bundles the harm and fault concepts, the burden of proof is assigned in the same fashion on each issue. As fair use is an affirmative defense, it is often believed that the burden of proving a fair use falls on the defendant. As a result, once the plaintiff has proved copying and substantial similarity, the defendant is burdened with proving that either this was a case of no harm or a case of no fault. However, as this Subsection argues, there is a reason to believe that the burden of proof should be assigned differently on these two issues.

a) The Theory of Burden Shifting

From a utilitarian viewpoint, the burden of proving harm and fault typically is placed on the plaintiff for reasons of minimizing the administrative costs of the litigation process.²²⁵ According to the conventional view, if the plaintiff is not required to prove the existence of harm and fault he will be given an incentive to begin cases that are not

harm to the plaintiff is negated by proof that the defendant's actions also benefited the plaintiff. For example, consider the case in which A negligently crashes his car into B, causing B to break his leg. Further imagine that, had A not crashed into B, drunk driver C would have crashed into B later, causing total paralysis. A's actions have negatively affected B, but arguably caused a net positive because he saved B from even greater harm. However, this does not exculpate A from liability. The fact that A caused the legally cognizable harm is sufficient to hold him liable, despite the arguable case that his actions may have benefited B. Likewise, in Google Books, the fact that the service may cause market substitution is sufficient to find the right has been infringed. This is no less so because the service also produces benefits for the copyright holder. Furthermore, it is highly questionable that Google Books will affect every copyright holder in the same way. While it may increase sales for some authors, it seems equally plausible that it will reduce sales for other authors. Finally, the court's use of empirical evidence is questionable. In *Perfect 10*, the court argued that there was no empirical proof of harm caused by Google, in this case, the court gives the benefit of the doubt to Google, despite no empirical proof that the service actually benefited authors. It seems somewhat unfair to require empirical proof of harm before holding Google liable, but then not to require empirical evidence from Google to substantiate their claim that they actually benefit the copyright holders. This article takes the view that the courts in these cases have done themselves no favors. They have reached the correct conclusion, but in a fashion that made their own jobs harder and has prejudiced future litigants. The court has improperly tried to use the language of "no harm" to justify their conclusion. They could have avoided using convoluted reasoning by accepting that the defendant's actions more than likely harmed some copyright holders, but concluding that, nonetheless, such harm was justifiable and thus not wrongful. These should have been "no fault" cases.

225. Bruce L. Hay & Kathryn E. Spier, *Burdens of Proof in Civil Litigation: An Economic Perspective*, 26 J. LEGAL STUD. 413, 413 (1997) ("Our principle claim is that courts can use the burden of proof to limit the costs of resolving a dispute.")

meritorious.²²⁶ By requiring the plaintiff to introduce evidence of harm and fault, we ensure that the plaintiff only brings cases that are likely to succeed, and thus reduce courts' expenditure on meritless litigation. Thus, the normal assignment of burdens results in "economizing on the time of the tribunal."²²⁷

Of course, there are exceptions to this standard rule. As Richard Posner points out, saying that placing the burden of proving a particular element of the case on the plaintiff reduces administrative costs assumes that the cost to the plaintiff of gathering the evidence to prove his point is no greater than the cost to the defendant of obtaining contrary evidence.²²⁸ In cases such as *res ipsa loquitur*, the burden of proof is shifted onto the defendant because it is easier, and therefore cheaper, for the defendant to prove the absence of fault than for the plaintiff to prove the existence of it.²²⁹

b) The Burden of Proving Harm and Fault

The question we are presented with is: how can we assign the burden of proof in order to reduce the administrative cost of the copyright system? But at this point, we are faced with the fact that the fair use analysis is composed of two different inquiries: firstly whether legally cognizable harm existed, and secondly whether fault existed. This Article takes the view that reducing administrative costs requires the burden to be assigned differently on these two distinct issues.

Turning first to the question of legally cognizable harm, it seems highly plausible that, in most cases, the plaintiff is far better suited to prove the existence of market harm than the defendant is to prove the lack of market harm. Not only does the plaintiff already have the most relevant information regarding his expected market and the loss in sales attributable to the defendant's copying, but requiring the defendant to prove the absence of market harm requires the proof of a negative. Given the complexity involved in proving a negative, this situation would seem to be much more costly than asking the plaintiff to prove the positive existence of market harm.

However, the same cannot be said on the issue of fault. Once market harm is proved, it then falls to the court to determine whether the copying

226. Bruce L. Hay, *Allocating the Burden of Proof*, 72 IND. L.J. 651, 656 (1997) ("The plaintiff, being the one pressing for judicial intervention, should therefore be required to show that she is entitled to the relief she seeks. Such a rule ensures that the legal system will—in general—only intervene in cases where there is a good reason (where relief is warranted).").

227. POSNER, *supra* note 86, at 646–47.

228. *Id.*

229. *Id.*

was in the public interest. This requires evidence that the copying will produce benefits in terms of access that is greater than the cost of reduced authorial incentives represented by the plaintiff's market harm. It would appear that the defendant is better placed to introduce such evidence. The defendant is the one using the copyrighted work in these instances. As the person most familiar with the purpose and character of the use, he seems better placed to demonstrate why this use is in the broader public interest. For example, if we reconsider *Perfect 10 v. Amazon*, while Perfect 10 is in the best position to prove that Google's use caused harm to their cell-phone download market, Google is ideally suited to explain why their product is so socially valuable that its access-benefits outweigh its incentive-costs.²³⁰

3. *Solution: Restructuring the Fair Use Analysis*

The most elegant solution would be to separate the concepts of harm and fault into their own doctrines. One plausible way to do this would be to integrate the harm inquiry into the outcome inquiry. That is, we would remove the market harm question from the fair use doctrine, and reinsert it into the substantial similarity doctrine. One could envision a regime in which the court uses the concept of market substitution in order to determine whether a substantially similar work exists. That is, a defendant's work will only be considered substantially similar if consumers view the two works as substitutes for one another in the market. In doing so, the burden of proving market substitution, and therefore that a rights-infringement has occurred, would fall squarely on the plaintiff.

The reality is, however, that such a drastic change is very unlikely to occur. The market harm question is bound into the fair use analysis by statute. Courts are unlikely to break up and segregate this congressionally mandated doctrine. Furthermore, post the enactment of the 1976 Act, courts have decided several hundred fair use cases.²³¹ In all of them, the question of harm has been an integral part of the fair use analysis. Changing direction now would be an unprecedented step to say the least.

Luckily however, such severe formal restructuring is not necessary. Instead, we could potentially restructure the fair use analysis in such a way as to facilitate the required conceptual separation. While Congress has indicated

230. In this respect, this article departs slightly from the prescriptive recommendation offered by Professor Hetcher. Professor Hetcher has suggested that the entire fair use doctrine ought to become part of the prima facie case. This would require the plaintiff to prove that the defendant's use was unfair before the defendant's prima facie liability is established. See Hetcher, *Fault*, *supra* note 136. Also on this topic, see Christopher Jon Sprigman, *Copyright and the Rule of Reason*, 7 J. TELECOMM. & HIGH TECH. L. 317 (2009).

231. See generally Samuelson, *supra* note 107.

that market harm is an element of the fair use analysis, there is no legislative requirements restricting how courts structure and apply the fair use analysis. The doctrine was left open-ended so that courts would have the ability to continue to shape it into a workable tool.²³² For example, the judiciary has the substantial interpretive room in defining the content of each factor; has the authority to add on additional factors; has the ability to weight the factors differently in different cases; can sequence their discussion of the factors creatively; and may decide how to assign the burden of proof. This raises the possibility that we can restructure the analysis that takes place under the heading of fair use in a way that puts some conceptual clear water between the issues of harm and fault.

Therefore, this Article suggests that copyright infringement analysis adopt the following structure. After the plaintiff proves copying (conduct) and substantial similarity (outcome), the defendant will be held liable as a *prima facie* matter. The defendant may then claim fair use. At which point, the burden should lie on the plaintiff to demonstrate that he suffered some market harm. If he fails to provide evidence that, on the balance of probabilities, establishes market harm, then he has not proved that his right has been infringed. The court then dismisses the case. Alternatively, if he does prove market harm, the burden ought to shift to the defendant to prove why there was no fault. That is, the defendant must prove, again on the balance of probabilities, that this use would lead to greater benefits in terms of access than cost in lost future incentives. Much of this argument would occur under the transformative doctrine. In cases like *Perfect 10*, *Bill Graham Archives*, and *Google Books*, the defendant would demonstrate that he has transformed the meaning in such a way that leads to great new social value. If the defendant successfully demonstrates that this was a case of “no fault,” then the case is dismissed. Alternatively, if he fails to demonstrate no fault, the court holds him liable and proceeds to the question of remedies. In this way, not only are judges guided towards treating harm and fault distinctly, but the burden of proof on each of these issues is assigned within the fair use analysis in a way that shall minimize administrative costs.

232. According to the legislative history behind the Copyright Act 1976, the wording of § 107 was “intended to restate the present judicial doctrine of fair use, not to change, narrow, or enlarge it in any way. . . . Beyond a very broad statutory explanation of what fair use is and some of the criteria applicable to it, the courts must be free to adapt the doctrine to particular situations on a case-by-case basis.” See H.R. Rep. No. 94-1476, at 66 (1976).

V. CONCLUSION

Despite the widespread and orthodox belief that copyright infringement is a strict liability tort, the reality is far more complex. This Article has advanced the thesis that copyright infringement is not a strict liability tort, but a fault-based tort. Liability is conditioned upon four elements: conduct (copying), outcome (substantial similarity), harm (market harm) and fault (unfairness). The fundamental question in the fair use analysis is a fault question. Like the issue of reasonableness in negligence, liability is only imposed upon those who fail to live up to a standard set by society for the purpose of maximizing social welfare. Only those who fail to reach the standard, and thus wrongfully force negative consequences on others in society, are held liable. Sadly, however, currently the fair use analysis is not only composed of a fault inquiry but also includes a harm inquiry. The question of whether the defendant has caused harm, and thus infringed a right, falls away from the *prima facie* case, and slips into the fair use analysis, resulting in much confusion. It is unlikely that this state of affairs will change any time soon. Therefore, this Article has proposed a restructuring of the fair use analysis. While not an ideal solution, this will at least mitigate the negative consequences created by the current collapsing of the harm and fault inquiries into the same doctrine.

Although these arguments are directed primarily at intellectual property scholars, this Article also holds lessons for tort theoreticians. Ascertaining exactly what makes negligent conduct a type of fault is a complex question that has long plagued tort scholars. This difficulty has occurred along both moral and legal dimensions. Firstly, there is the question of why harm caused without any bad will on the part of the defendant is morally wrong. Secondly, as the legal concept of fault may or may not be exactly the same as the moral concept of fault,²³³ the question emerges, what is the legal concept of fault and why does negligent behavior fit within that definition? At the moment, the best analytic theory states that failing to comply with a standard is a type of fault in law, and hence, negligence is faulty conduct. However, if failing to live up to a standard of conduct is a type of fault, then why is failing to live up to the requirement that people only copy fairly any less a type of fault? Thus, copyright infringement provides a test case and a challenge for our definitions of fault and negligence. This Article presents the view that failing to comply with a standard of conduct is a form of fault and therefore copyright infringement is a fault-based tort. If some cannot accept that

233. See generally H.L.A. Hart, *Positivism and the Separation of Law and Morals*, 71 HARV. L. REV. 593 (1958).

conclusion, then this Article calls upon those scholars to produce a clearer definition of fault—one that is simultaneously capable of demonstrating why negligent conduct is fault, but unfair conduct is not.

Finally, a last word must be made about international copyright. This Article's discussion of the "fairness" liability rule in copyright has been restricted to the United States, which relies heavily on the fair use doctrine. Yet, as scholars of international copyright law will accurately point out, most countries do not adopt a fair use doctrine. In these jurisdictions, copyright infringement is arguably still a strict liability tort. They impose liability on the basis of copying and substantial similarity, without regard to either the defendant's mental state or his conformity with a standard of conduct. However, it is interesting to note that in recent years the fair use doctrine has grown internationally. A number of countries, such as South Korea,²³⁴ the Philippines,²³⁵ and Singapore,²³⁶ have adopted the standard. Some countries, such as Canada, have amended their existing exceptions to copyright infringement to become more fair use-like in character.²³⁷ Meanwhile, other countries, such as the United Kingdom,²³⁸ Australia,²³⁹ and Ireland,²⁴⁰ are seriously considering adopting the doctrine. In the discussions taking place in these jurisdictions, there is a recurrent belief that adopting fair use will provide the necessary incentives for authors and copyists to create and use copyrighted works in ways that will generate economic growth in the so-called "digital economy." This author interprets the internationalization of fair use as the rejection of strict liability in favor of the more efficient fault liability rule that the fair use doctrine instantiates. However, the exact

234. Byeon ri sa beob il bu gae jeong beop ryul [On Partial Amendment to the Patent Attorney Act], Act No. 11962, July 30, 2013 (S. Kor.).

235. An Act Amending Certain Provisions of Republic Act No. 8293, Otherwise Known as the "Intellectual Property Code of the Philippines", and for Other Purposes, Rep. Act No. 10372, § 12 (Mar. 22, 2013) (Phil.).

236. Copyright Act, Ch. 63, § 35 (2006) (Sing.).

237. THE COPYRIGHT PENTALOGY: HOW THE SUPREME COURT OF CANADA SHOOK THE FOUNDATIONS OF CANADIAN COPYRIGHT 157–81 (Michael Geist ed., 2013).

238. IAN HARGREAVES, DIGITAL OPPORTUNITY: A REVIEW OF INTELLECTUAL PROPERTY AND GROWTH 1 (2011); For an understanding of the U.K.'s current approach to copyright infringement and exceptions, see William Cornish & Graeme Dinwoodie, *Intellectual Property*, in ENGLISH PRIVATE LAW §§ 6.64–6.69 (Andrew Burrows ed., 3d ed. 2013).

239. AUSTRALIAN LAW REFORM COMM'N, COPYRIGHT AND THE DIGITAL ECONOMY 3 (2014).

240. GOV'T OF IR., COPYRIGHT REVIEW COMM., COPYRIGHT AND INNOVATION: A CONSULTATION PAPER 89 (2012), available at http://www.djei.ie/science/ipr/crc_consultation_paper.pdf.

motivation and significance of this global shift is the subject of another article.²⁴¹

241. Patrick R. Goold, *The Fair Use Revolution in Global Copyright Law* (unpublished manuscript, on file with the author).