BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 24 NUMBER 2 SPRING 2009

TABLE OF CONTENTS

ARTICLES

SUBSCRIBER INFORMATION

The *Berkeley Technology Law Journal* (ISSN 1086-3818), a continuation of the *High Technology Law Journal* effective Volume 11, is edited by the students of the University of California School of Law, Berkeley (Boalt Hall), and published four times each year (March, June, September, January) by the Regents of the University of California, Berkeley, Journal Publications, School of Law, 311 Boalt Hall, University of California, Berkeley, CA 94720-7200. Periodicals Postage Rate Paid at Berkeley, CA 94704-9998, and at additional mailing offices. POSTMASTER: Send address changes to Journal Publications, 311 U.C. Berkeley School of Law, University of California, Berkeley, CA 94720-7200.

Correspondence. Address all correspondence regarding subscriptions, address changes, claims for non-receipt, single copies, advertising, and permission to reprint to Journal Publications, 311 Boalt Hall, School of Law, Berkeley, CA 94720-7200; (510) 643-6600; JournalPublications@ law.berkeley.edu. Authors: see section entitled Information for Authors.

Subscriptions. Annual subscriptions are $65.00 for individuals, and $85.00 for organizations. Single issues are $27.00. 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 ($27.00) will be charged for replacement. Overseas delivery is not guaranteed.

Form. The text and citations in the *Journal* conform generally to the UNITED STATES GOVERNMENT PRINTING OFFICE STYLE MANUAL (29th ed. 2000) and to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 18th ed. 2005). Please cite this issue of the *Berkeley Technology Law Journal* as 24 BERKELEY TECH. L.J. ___ (2009).

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*, selected materials related to technology law, and links to other related pages. Author, volume, and subject indexes may also be found in Volume 20, Number 4 (2005) of the *Journal*.

MODERNIZING PATENT LAW'S INEQUITABLE CONDUCT DOCTRINE

By Christopher A. Cotropia†

TABLE OF CONTENTS

---

† Professor of Law, Intellectual Property Institute, University of Richmond School of Law. I would like to thank Dawn-Marie Bey, Jim Gibson, Mary Heen, Corinna Lain, Shari Motro, Michael Risch, Sean Seymore, and the participants at the 2007 Works in Progress Intellectual Property Colloquium at the American University Washington College of Law and a Richmond Law School Colloquium for their comments on an earlier draft of this Article.

I. INTRODUCTION

The inequitable conduct doctrine governs a patent applicant's duties before the United States Patent and Trademark Office ("USPTO"). The doctrine requires the inventor to disclose information to the USPTO that is relevant to the patentability—the utility, novelty, nonobviousness, and adequate disclosure—of the invention at issue.[1] The doctrine is pervasive, imposing a duty to disclose and be truthful in every correspondence with the USPTO.

The penalty for failing to discharge this duty is dramatic—coined an "atomic bomb" by one Federal Circuit judge.[2] A finding of inequitable conduct renders the entire patent unenforceable for the rest of the patent term, even when the undisclosed information was material to only a particular patent claim.[3] In some cases, the doctrine extends its reach to related patents, rendering them unenforceable as well.[4] The resulting unenforceability persists even if the invention actually meets the patent requirements.[5] The nature of the inequitable conduct doctrine makes it unique in patent law, in that it is an individual's failure to disclose—rather than an inherent trait of the claimed invention—that results in the denial of protection for the invention and other related patents.

Given the all-encompassing nature of the inequitable conduct doctrine and its death-penalty-like remedy, it is not surprising that the doctrine has garnered much attention and criticism since its inception. The Federal Circuit has gone out of its way on more than one occasion to criticize aspects of the doctrine.[6] There has been a tremendous amount of recent activity at

---

1. *See* Molins PLC v. Textron, Inc., 48 F.3d 1172, 1178 (Fed. Cir. 1995) (setting forth the three basic elements of inequitable conduct—materiality, non-disclosure, and intent); 37 C.F.R. § 1.56 (2008) (describing the type of information a patent applicant is under a duty to disclose).

2. *See* Aventis Pharma S.A. v. Amphastar Pharms., Inc., 525 F.3d 1334, 1349 (Fed. Cir. 2008) (Rader, J., dissenting).

3. *See* Kingsdown Med. Consultants, Ltd. v. Hollister, Inc., 863 F.2d 867, 877 (Fed. Cir. 1988) (en banc) ("When a court has finally determined that inequitable conduct occurred in relation to one or more claims during prosecution of the patent application, the entire patent is rendered unenforceable.").

4. If there is a pattern of inequitable conduct, unenforceability can transfer from one patent to another. *See* Consol. Aluminum Corp. v. Foseco Int'l, Ltd., 910 F.2d 804, 812 (Fed. Cir. 1990).

5. *See* Star Scientific, Inc. v. R.J. Reynolds Tobacco Co., 537 F.3d 1357, 1365 (Fed. Cir. 2008) (noting that "the penalty for inequitable conduct is so severe, the loss of the entire patent even where every claim clearly meets every requirement of patentability").

6. Judges on the court have characterized the doctrine as "overplayed," *Kimberly-Clark Corp. v. Johnson & Johnson*, 745 F.2d 1437, 1454 (Fed. Cir. 1984), and labeled its

the Federal Circuit regarding the proper strength of the doctrine, resulting in enough of a conflict to prompt two judges over a five-month period to draft dissents addressing the doctrine in general.[7] One Federal Circuit judge recently expressly called for reconsideration of the doctrine by the whole court.[8] Patent practitioners constantly monitor and critique the development of the doctrine, partly because it focuses on "the person rather than the patent."[9] This attention by both judiciary and bar spurred two major patent system studies to discuss possible modifications to the inequitable conduct doctrine.[10] One study went so far as to suggest the elimination of the doctrine altogether.[11] Congress has also begun to pay attention, with essentially every draft of its recent patent reform legislation containing some amendment to the doctrine.[12] Even the USPTO has suggested rule changes that would affect the doctrine.[13]

Every facet of the patent system—Congress, the Federal Circuit, the USPTO, and patent practitioners—is concerned about the state of the inequitable conduct doctrine. No consensus exists, however, as to what is wrong with the doctrine and how it should be changed. As a result, the dis-

---

habitual assertion in litigation as "an absolute plague," *Burlington Indus., Inc. v. Dayco Corp.*, 849 F.2d 1418, 1422 (Fed. Cir. 1988).

7. *See* Praxair, Inc v. ATMI, Inc., 543 F.3d 1306, 1329–31 (Fed. Cir. 2008) (Lourie, J., dissenting); *Aventis*, 525 F.3d at 1349 (Rader, J., dissenting).

8. *See* Larson Mfg. Co. of S. Dakota, Inc. v. Aluminart Prods. Ltd., 559 F.3d 1317, 1342, 1344 (Fed. Cir. Mar. 18, 2009) (Linn., J., concurring) (concluding that the recent inequitable conduct precedent "has significantly diverged from the Supreme Court's treatment of inequitable conduct" and "the time has come for the court to review the issue en banc").

9. 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).

10. *See* Fed. Trade Comm'n, To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy 12–13 (2003), *available at* http://www.ftc.gov/os/2003/ 10/innovationrpt.pdf; Nat'l Research Council, A Patent System for the 21st Century 121–23 (Stephen A. Merrill et al. eds., 2004).

11. *See* NAT'L RESEARCH COUNCIL, *supra* note 10, at 123.

12. *See, e.g.*, Patent Reform Act of 2007, H.R. 1908, 110th Cong. § 123 (2007); Patent Reform Act of 2006, S. 3818, 109th Cong. § 5 (2006); Patent Reform Act of 2005, H.R. 2795, 109th Cong. § 136 (2005).

The 2009 legislative draft pending before the Senate does not contain any inequitable conduct language. *See* Patent Reform Act of 2009, S. 515, 111th Cong. (2009). Senator Orin Hatch, however, is still pushing for inequitable conduct reform to be included in the legislation. *See* 155 CONG. REC. S2715 (daily ed. Mar. 3, 2009) (statement of S. Hatch).

13. *See* Changes To Information Disclosure Statement Requirements and Other Related Matters, 71 Fed. Reg. 38,808 (July 10, 2006) (to be codified at 37 C.F.R. pt. 1) (proposing to change the IDS requirements to include relevancy statements, in addition to other requirements).

jointed discussion has created inconsistency at the legislative level—with reform proposals driven by whichever critique Congress is focused on in a given session.

Running in parallel with the conversation about the inequitable conduct doctrine is a broader discussion of patent quality. One of the major focuses of the patent reform movement is to ensure that the USPTO issues only those patents that claim truly patentable inventions.[14] The number of patent applications is rising exponentially, the time a patent examiner can spend examining them is decreasing, and the quality of patent applications and the information available for examination is dropping.[15] All of these factors cause more patents to issue that, in actuality, should not be issued. The resulting "bad" patents—patents that fail to meet the patentability requirements—are harmful, creating detrimental societal costs via hold-ups and the *in terrorem* effects that invalid patents create.[16]

These two topics—the inequitable conduct doctrine and patent reform in general—are being addressed with increasing frequency by academics. Academics have written articles on the patent examination process, ways to reform it, and the negative impact of issuing bad patents.[17] Likewise, many scholars have written articles specifically on the inequitable conduct doctrine.[18] However, this scholarship has failed to fully link these two areas. As a result, it has failed to engage in two basic, interrelated exercises that would greatly assist the discourse on the inequitable conduct doctrine and patent reform in general.

First, no one has attempted a comprehensive, theoretical analysis of how the inequitable conduct doctrine as a whole affects patent applicants,

---

14. *See* 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, 270, 276–78 (2007) ("In the last few years, widespread dissatisfaction with the patent system--and particularly with the perceived poor quality of issued patents--has spurred a broad range of groups to call for reform.").

15. *See* Doug Lichtman & Mark A. Lemley, *Rethinking Patent Law's Presumption of Validity*, 60 STAN. L. REV. 45, 46–47 (2007) (identifying these problems as the cause of the USPTO's "mistakes" while reviewing patent applications).

16. *See* Christopher R. Leslie, *The Anticompetitive Effects of Unenforced Invalid Patents*, 91 MINN. L. REV. 101, 113 (2006) ("[S]ome invalid patents can deter market entry and decrease consumer welfare even without active enforcement.").

17. *See, e.g.*, John R. Thomas, *Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties*, 2001 U. ILL. L. REV. 305, 305 (2001) ("By awarding prior art informants with a bounty assessed against applicants, the Patent Office can restore order to the patent system and reduce its social costs.").

18. *See, e.g.*, Robert J. Goldman, *Evolution of the Inequitable Conduct Defense in Patent Litigation*, 7 HARV. J.L. & TECH. 37 (1993) (tracing "the development of the law concerning inequitable conduct from the beginning of the patent system").

patent examination, and potential inventors.[19] Performing a fundamental analysis of the doctrine would provide a framework by which proposed changes could be tested. Such an analysis would also flesh out how the doctrine plays a significant role in the patent system overall.

Second, and related, almost all of the scholarship on the inequitable conduct doctrine has kept the discussion tied to its equitable roots, focusing on the doctrine as an ethical tool.[20] These discussions are of little help when the specific criticisms of the doctrine are not morally focused. Furthermore, any proposed change to the doctrine needs to be considered in the context of broader reforms rooted in the utilitarian theory that underlies American intellectual property law. Thinking on the inequitable conduct doctrine needs to be modernized and framed in the same utilitarian terms that form the foundation for the patent system. This Article attempts to fill these scholarly holes and answer the question as to how the doctrine should be changed and used to improve the patent system.

This Article's main finding is that the inequitable conduct doctrine has the ability to improve patent quality as long as the inherent tendency to overcomply with the doctrine by overloading the USPTO with information is kept in check. The Article reaches this conclusion by proceeding in five parts. Part II describes the current thinking on the inequitable conduct doctrine, with particular focus on the major critiques of the doctrine and proposed legislative and administrative responses. Part III of the Article begins the construction of a fundamental, conceptual framework for the doctrine by explaining how it impacts both patent quality and patent examination. If properly calibrated, the doctrine can improve both the quality of the patent application (by increasing the patent attorney's knowledge and

---

19. A majority of the articles are sound, but focus on specific parts of the doctrine, particular proposed statutory changes and individual Federal Circuit cases. *See, e.g.*, David Hricik, *Where the Bodies Are: Current Exemplars of Inequitable Conduct and How to Avoid Them*, 12 TEX. INTELL. PROP. L.J. 287, 289 (2004) (discussing, in detail, the various fact patterns that have supported a finding of inequitable conduct); James Cronin, Comment, *Inequitable Conduct and the Standard of Materiality: Why the Federal Circuit Should Use the Reasonable Patent Examiner Standard*, 50 ST. LOUIS U. L.J. 1327 (2006) (discussing the materiality requirement of the inequitable conduct doctrine).

20. One commentator even affirmatively dismissed the linkage between the doctrine and patent quality, concluding that reforms should focus on "punishing bad behavior on the part of applicants." Cronin, *supra* note 19, at 1360. Articles have mentioned the doctrine's possible impact on patent quality. *See, e.g.*, Kevin Mack, Note, *Reforming Inequitable Conduct to Improve Patent Quality: Cleansing Unclean Hands*, 21 BERKELEY TECH. L.J. 147, 166–69 (2006). These articles mention this linkage only in passing, failing to fully develop the discussion conceptually. For example, no article has discussed the huge potential for over compensation under the doctrine, a significant aspect of the doctrine's impact on patent quality developed in this Article. *See infra* Parts IV, V.

care) and the quality of the examination (by acting as an information producer and verifier).

The doctrine's potential impact is not all positive. The tremendous incentive for applicants to overcomply can actually decrease patent quality. Part IV completes the conceptual framework by detailing how the doctrine, through the extreme legal and extra-legal costs it currently imposes, incentivizes inventors and, in particular, patent attorneys to overcomply by submitting all information, regardless of relevance, to the USPTO. Part V explains how this overcompliance negatively affects patent examination and the patent system by causing information overload that hampers the USPTO's ability to operate effectively and by creating high compliance costs that price inventors out of the patent system. Finally, in Part VI, the Article uses this framework to suggest changes that maintain the positive effects of the doctrine on patent quality while minimizing overcompliance.

II. CURRENT THINKING ON THE INEQUITABLE CONDUCT DOCTRINE

Before a conceptual framework regarding the inequitable conduct doctrine can be developed, the basic process of obtaining a patent, the current contours of the inequitable conduct doctrine, and the debate surrounding it need to be detailed. This Part begins with a description of the patent prosecution process. It then details the three requirements of the inequitable conduct doctrine—materiality, disclosure, and intent—and available remedies. This Part concludes by describing the popular critiques of the doctrine, the proposed legislative responses, and the major disconnects in this discourse. Those familiar with the patenting process and the inequitable conduct doctrine can skip directly to Section II.C.

A. Basics of Patent Prosecution

To patent an invention in the United States, an inventor must file a patent application with the USPTO. The patent application contains a textual and graphical description of the invention called the specification. The specification includes general statements regarding the technical background of the invention, the problem it is trying to solve, and some specific examples—embodiments—of the invention.[21] The patent application

21. *See* 35 U.S.C. § 112, para. 1 (2006) (reciting the written description, enablement, and best mode requirements for the specification); Christopher A. Cotropia, *Patent Claim Interpretation and Information Costs*, 9 LEWIS & CLARK L. REV. 57, 68–69 (2005) (discussing the various kinds of information patent law requires the inventor put in the patent's specification); 37 C.F.R. §§ 1.71, 1.77(b) (2008) (indicating the various types of information the specification should include).

also includes a set of claims. Each claim is a single sentence that defines the exact invention the inventor wishes to protect.[22] The application can either be filed by the inventor herself or by a patent attorney or agent, who must be a member of the USPTO bar and represents the inventor.[23]

Once filed, the application is given to a patent examiner assigned to the invention's technological area. The examiner reviews each claim to determine whether it meets the requirements for patentability—whether the claim defines an invention that is useful, novel, and nonobvious, and whether the specification adequately describes and enables the claimed invention.[24] To make this determination, the examiner must first gain an understanding of the exact scope of the claims. From there, the examiner searches for information—referred to as "prior art"—that might render the claims invalid.[25] This information can come in many forms, such as scien-

---

22. *See* 35 U.S.C. § 112, para. 2; Aro Mfg. Co. v. Convertible Top Replacement Co., 365 U.S. 336, 339 (1961) ("[T]he claims made in the patent are the sole measure of the grant . . . ."); Christopher A. Cotropia, *Patent Claim Interpretation Methodologies and Their Claim Scope Paradigms*, 47 WM. & MARY L. REV. 49, 61–62 (2005).

23. 37 C.F.R. § 1.31 (2008). For the purposes of this Article, patent attorneys and patent agents, those who are members of the USPTO bar but not lawyers, are referred to collectively as "patent attorneys." Most patent applicants are represented by patent attorneys or patent agents—few go "pro se." *See* Martin B. Schwimmer, *Domain Names and Everything Else: Trademark Issues in Cyberspace*, *in* UNDERSTANDING BASIC TRADEMARK LAW 1998, at 263, 268 (PLI Patents, Copyrights, Trademarks, & Literary Prop., Course Handbook Series No. G0-001S, 1998) ("Finally, the 'constituency' of the PTO, namely patent and trademark applicants, are represented, for the most part, by attorneys who, as members of bar committees, communicate with the PTO.").

24. 35 U.S.C. §§ 101–103, 112 (2006) (setting forth the requirements for patentability); 37 C.F.R. § 1.104(c) (2008).

The USPTO is currently forced to examine applications under extreme time constraints and in the face of a significant backlog of pending applications. *See* Eric B. Chen, *Conflicting Objectives: The Patent Office's Quality Review Initiative and the Examiner Count System*, 10 N.C. J. L. & TECH. 28, 30 (2009) ("Despite this success, several challenges continue to plague the USPTO, namely the backlog of unexamined patent applications, concerns over examiner attrition, and the increasing volume of continuing applications. . . ."). This situation leads to long delays between the filing of a patent application and an examiner's action on that application. *See* Gary C. Ganzi, *Patent Continuation Practice and Public Notice: Can They Coexist?*, 89 J. PAT. & TRADEMARK OFF. SOC'Y 545, 563 (2007) ("However, with present application pendency delays, a typical patent application does not usually even receive the benefit of a first examiner office action at the USPTO within the publication time of eighteen months after filing."). These circumstances have import for later analysis in this Article. *See infra* Section III.A.

25. Information relevant to examination, while described in all subsections of 35 U.S.C. § 102 (2006), falls into two basic categories, both defining the universe of "prior art." The first set of prior art is that information produced by those other than the inventor prior to the date of the invention. *See, e.g.*, § 102(a). The second is art produced by any-

tific articles, general publications, other United States patents or patents issued by other countries, and general public knowledge or use.[26]

The applicant can submit potential prior art to the USPTO. Such art may be cited in the background section of the patent application to give context to the claimed invention.[27] The information is usually submitted via an information disclosure statement ("IDS").[28] Depending on the timing of the submission of the IDS, the applicant may have to request another round of examination to give the examiner time to consider the submitted information.[29]

The examiner compares the prior art to the application's claims to determine whether the claims are novel and nonobvious.[30] The examiner also looks to see if the specification adequately describes and enables the claimed invention.[31] Based on the results of this initial examination, the examiner issues an "office action" to the applicant describing the examiner's findings and identifying which claims she believes to be and not to be patentable and the reasons for this conclusion.[32]

---

one—including the inventor—more than one year before the patent application's filing. *See, e.g.*, § 102(b).

26. *See, e.g.*, §102(a)–(b) (detailing these different types of prior art).

27. 37 C.F.R. § 1.77(e) (2008).

28. 37 C.F.R. §§ 1.97, 1.98 (2008) (detailing the filing procedure of an IDS and its content). Prior art may also be disclosed in the patent specification itself. *See, e.g.*, 37 C.F.R. § 1.71(b) (2008) (indicating that applicant should distinguish her invention "from what is old").

29. *See* 37 C.F.R. § 1.98(b)–(d) (listing the timing requirements for filing a proper IDS); 35 U.S.C. § 132(b) (2006) (establishing a method for continuing examination—a request for continued examination ("RCE")).

30. The examiner looks to see if the claims have already been disclosed in the prior art—that is, whether they are not novel or statutorily barred. *See* 35 U.S.C. § 102; *In re* Paulsen, 30 F.3d 1475, 1478–79 (Fed. Cir. 1994) ("A rejection for anticipation under section 102 requires that each and every limitation of the claimed invention be disclosed in a single prior art reference."). The examiner also looks to see if pieces of the prior art would have been combined together to duplicate an application's claim—that is, whether the claim is obvious. *See* 35 U.S.C. § 103 (2006); KSR Int'l Co. v. Teleflex, Inc., 550 U.S. 398, 405 (2007). The Court in *KSR* stated:

> Section 103 forbids issuance of a patent when 'the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.'

Id.

31. 35 U.S.C. § 112, para. 1 (2006); *In re* Fisher, 421 F.3d 1365, 1378 (Fed. Cir. 2005).

32. 35 U.S.C. § 151 (2006) (detailing the issuance of a patent).

The applicant responds to the office action by rebutting the examiner's analysis, amending the patent's claims to overcome the examiner's objections, or canceling patent claims altogether.[33] The examiner then reviews the applicant's response and either agrees and allows the claims or does not and maintains the rejections. The examination ends when either some of the patent claims are allowed or the patent is abandoned altogether. Once a patent is issued, it may be used to exclude others from practicing the claimed invention.[34] The patent claims that the USPTO concludes are patentable enjoy a strong presumption of meeting the patentability standards, requiring a challenger to prove by clear and convincing evidence that the claims are invalid.[35]

This whole process is secret, with the application and the correspondence only being publicly disclosed when and if the patent issues (or sooner if the applicant so elects).[36] The process is ex parte with the examiner representing the public's interest and is meant to be non-adversarial.[37] The public may comment on or submit art relating to a pending application,[38] but these options are rarely used.[39] This makes the patent examiner, the reviewing authorities within the USPTO, and the reviewing courts incredibly important to the patent process. They are essentially the only gate-

33. 35 U.S.C. § 132(a) (2006) (noting that examination continues when "the applicant persists in his claim for a patent, with or without amendment").

34. 35 U.S.C. § 271 (2006).

35. *See* 35 U.S.C. § 282 (2006); Am. Seating Co. v. USSC Group, Inc., 514 F.3d 1262, 1267 (Fed. Cir. 2008) (noting that the presumption of validity can be overcome with only "clear and convincing evidence"). This strong presumption makes examination incredibly important. *See* Lichtman & Lemley, *supra* note 15, at 47 (concluding that the presumption makes "issuance mistakes hard to reverse").

36. The application, and all related correspondence, is published eighteen months after filing, unless the applicant requests no such publication. 35 U.S.C. § 122(b) (2006).

37. *See* Hercules Inc. v. Exxon Corp., 434 F. Supp. 136, 152 (D. Del. 1977) ("The prosecution of an application before the Patent Office is not an adversary, but an ex parte proceeding."). The USPTO even views patent applicants as "customers." *See* ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS: HOW OUR BROKEN PATENT SYSTEM IS ENDANGERING INNOVATION AND PROGRESS, AND WHAT TO DO ABOUT IT 11, 20 (2004).

38. *See* 37 C.F.R. § 1.291 (2008) (describing the protest procedure and allowing prior art to be filed along with a protest).

39. Andrew Kopelman, Note, *Addressing Questionable Business Method Patents Prior to Issuance: A Two-Part Proposal*, 27 CARDOZO L. REV. 2391, 2413–14 (2006) (indicating that most do not use the protest mechanism because it is difficult to discover a pending application).

keepers to the initial and very important determination that an invention is worthy of twenty years of exclusivity.[40]

B. Requirements of the Inequitable Conduct Doctrine

The current inequitable conduct doctrine is judicially-created and impacts the entire patent prosecution process. The doctrine focuses on the patent application and related correspondence between the applicant and the USPTO during patent prosecution. It is comprised of three basic elements: materiality, intent, and disclosure. Specific remedies accompany a finding of inequitable conduct.

1. Materiality

The doctrine focuses on the disclosure of material information. Information is material if it is relevant to the patentability of the claimed invention being examined.[41] The standard for materiality is articulated in a USPTO regulation, 37 C.F.R. § 1.56 (Rule 56).[42] The most recent version of Rule 56 deems information material if

> (1) It establishes, by itself or in combination with other information, a prima facie case of unpatentability of a claim; or (2) It refutes, or is inconsistent with, a position the applicant takes in: (i) Opposing an argument of

40. *See* 35 U.S.C. § 131 (2006) (granting authority to the USPTO to examine applications and issue patents); 28 U.S.C. § 1295(a)(4)(A) (2006) (giving the Federal Circuit appellate jurisdiction over appeals from the Board of Appeals and Patent Interferences (BPAI)); 35 U.S.C. § 134 (2006) (authorizing the appeal from a final rejection to the BPAI). An applicant can, instead of appealing to the Federal Circuit, appeal a BPAI ruling to the United States District Court for the District of Columbia, whose decision is then appealed to the Federal Circuit. 35 U.S.C. § 145 (2006).

41. Dayco Prods., Inc. v. Total Containment, Inc., 329 F.3d 1358, 1363 (Fed. Cir. 2003) ("For many years this court held that materiality for purposes of an inequitable conduct determination required a showing that 'a reasonable examiner would have considered such prior art important in deciding whether to allow the parent application.'"). *Id.* at 1362 ("Information did not need to be prior art in order to be material, but 'instead embrace[d] any information that a reasonable examiner would substantially likely consider important in deciding whether to allow an application to issue as a patent.'") (quoting Akron Polymer Container Corp. v. Exxel Container, Inc., 148 F.3d 1380, 1382 (Fed. Cir. 1998)).

42. Purdue Pharma L.P. v. Endo Pharm. Inc., 438 F.3d 1123, 1129 (Fed. Cir. 2006) ("In evaluating materiality, this court has consistently referred to the standard set forth in PTO Rule 56."); 37 C.F.R. § 1.56(b) (2008). Rule 56 does not supplant the materiality standard articulated by the courts, it merely informs the standard. *See* Digital Control Inc. v. Charles Mach. Works, 437 F.3d 1309, 1314–16 (Fed. Cir. 2006).

> unpatentability relied on by the Office, or (ii) Asserting an argument of patentability.[43]

The materiality standard does not create a "but for" test, in which the information needs be disclosed only if it would actually render a pending patent claim invalid.[44] Instead, materiality is broader, requiring disclosure of information that would merely establish a prima facie case of invalidity that may be rebuttable.[45]

2. Intent

The non-disclosure of material information must be intentional to rise to the level of inequitable conduct.[46] An omission or misrepresentation of material information is considered intentional if the applicant actually intended to deceive or mislead the USPTO.[47] Gross negligence is not enough.[48] Circumstantial evidence can be used to prove the relevant party's intent.[49]

---

43. 37 C.F.R. § 1.56(b). This new standard, compared to the earlier "reasonable examiner" standard, "was not intended to constitute a significant substantive break with the previous standard." Hoffmann-La Roche, Inc. v. Promega Corp., 323 F.3d 1354, 1368 n.2 (Fed. Cir. 2003).

44. *See* Am. Hoist & Derrick Co. v. Sowa & Sons, Inc., 725 F.2d 1350, 1362–63 (Fed. Cir. 1984) (dismissing an objective "but for" test that would have required a prerequisite finding of invalidity to establish materiality).

45. *See* Monsanto Co. v. Bayer Bioscience N.V., 514 F.3d 1229, 1237 (Fed. Cir. 2008).

46. *See* Eli Lilly & Co. v. Zenith Goldline Pharm., Inc., 471 F.3d 1369, 1381 (Fed. Cir. 2006) ("[T]he trial court must also determine whether the evidence shows a threshold level of intent to mislead the PTO.").

47. *See* Ferring B.V. v. Barr Labs., Inc., 437 F.3d 1181, 1190 (Fed. Cir. 2006) ("Even if an omission is found to be material, the omission must also be found to have been made with the intent to deceive.").

48. *See* Kingsdown Med. Consultants, Ltd. v. Hollister Inc., 863 F.2d 867, 872 (Fed. Cir. 1988) (en banc). The Federal Circuit noted:

> We adopt the view that a finding that particular conduct amounts to 'gross negligence' does not of itself justify an inference of intent to deceive; the involved conduct, viewed in light of all the evidence, including evidence indicative of good faith, must indicate sufficient culpability to require a finding of intent to deceive.

Id. at 876.

49. *See* Hoffman-La Roche Inc. v. Lemmon Co., 906 F.2d 684, 688 (Fed. Cir. 1990). The use of circumstantial evidence is one of the main areas of recent disagreement amongst Federal Circuit judges. Intent needs to be independently proven. *See* Star Scientific, Inc. v. R.J. Reynolds Tobacco Co., 537 F.3d 1357, 1366 (Fed. Cir. 2008) ("[W]e have emphasized that 'materiality does not presume intent, which is a separate and essential component of inequitable conduct.'"). And many agree that finding direct evidence is

3. *Disclosure*

The doctrine requires the disclosure of material information.[50] A failure to disclose can occur in two instances—by omission or misrepresentation.[51] Omission is where the patentee fails to disclose material information in her filings with the USPTO. The typical non-disclosure by omission situation involves an applicant who fails to submit information which is in her possession, qualifies as prior art, and is material to one or more of the application's claims.[52] Misrepresentation, in contrast, occurs when the patentee does disclose information to the USPTO, but misrepresents a material aspect of the disclosed information.[53] Misrepresentation can also be more blatant, with the application simply submitting false in-

---

very unlikely. *Id.* Therefore, the question becomes what evidence is proper circumstantial evidence of intent, *id.*, and how does a court weigh such circumstantial evidence. *Compare id.* at 1366–67 (noting that a finding of intent based on circumstantial evidence must be "the single most reasonable inference"), *with* Praxair, Inc v. ATMI, Inc., 543 F.3d 1306, 1315 (2008) (stating that intent can be inferred if materiality is high, the applicant knew or should have known of the materiality, and the applicant failed to "come forward with any credible good faith explanation" for nondisclosure). This dispute over the standard for intent is important for the prescriptive portion of this Article. *See infra* Section VI.A.2.

50. 37 C.F.R. § 1.56(a) (2008). The regulation provides:

> Each individual associated with the filing and prosecution of a patent application has a duty of candor and good faith in dealing with the Office, which includes a duty to disclose to the Office all information known to that individual to be material to patentability as defined in this section.

Id.

51. *See* Pharmacia Corp. v. Par Pharm., Inc., 417 F.3d 1369, 1373 (Fed. Cir. 2005) ("[I]nequitable conduct includes affirmative misrepresentation of a material fact, failure to disclose material information, or submission of false material information, coupled with an intent to deceive." (quoting Molins PLC v. Textron, Inc., 48 F.3d 1172, 1178 (Fed. Cir. 1995))).

52. The failure to submit an earlier chemistry report and previous test data indicating that a prior canola oil formulation exhibited similar properties to the claimed canola oil formula is an example of material non-disclosure. Cargill, Inc. v. Canbra Foods, Ltd., 476 F.3d 1359, 1365–66 (Fed. Cir. 2007).

53. The facts of *Frazier v. Roessel Cine Photo Tech, Inc.*, 417 F.3d 1230 (Fed. Cir. 2005), provide a good example of a material misrepresentation. The patent at issue in *Frazier* claimed a "Z lens" that provided for an increased depth of field—allowing both a close-up object and a distance background to both appear in focus at the same time. *Id.* at 1234–36. During prosecution, the applicant submitted a videotape to help demonstrate the superiority of the invention over the prior art. *Id.* This submission was deemed a material misrepresentation because some of the most striking examples of depth of field in the video were from the unlabeled use of a "L-shaped lens," not the claimed Z lens. *Id.*

formation to the USPTO.[54] Again, there is disclosure, but the truth is not disclosed.

The inequitable conduct doctrine does not require the disclosure of information that is material if it is cumulative in light of information already provided to the USPTO.[55] Cumulative information is information already before the USPTO, albeit from a different source. Thus, not providing the USPTO with cumulative information is not, in fact, a failure to disclose.[56]

The duty to disclose imposed by the doctrine applies to more than just the inventor. Rule 56 extends the duty to all "[i]ndividuals associated with the filing or prosecution of [the] patent application."[57] This includes the attorney or agent who prepares or prosecutes the patent application.[58] The duty also applies to those who are "substantively involved" and associated with the inventor or her employer.[59]

The duty to disclose does not currently include a duty to search. The inequitable conduct doctrine only requires the applicant to disclose material information within her possession or the possession of those other individuals associated with prosecution. The duty does not, however, compel an applicant to actively search for additional prior art and, in turn, disclose it to the USPTO.[60]

---

54. *See* Digital Control, Inc. v. Charles Mach. Works, 437 F.3d 1309, 1316 (Fed. Cir. 2006) (concerning potentially false statements by the inventor to the USPTO about the use of a prior art boring tool).

55. Honeywell Int'l Inc. v. Universal Avionics Sys. Corp., 488 F.3d 982, 1000 (Fed. Cir. 2007) ("Information cumulative of other information already before the Patent Office is not material."); 37 C.F.R. § 1.56(b) (2008) ("[I]nformation is material to patentability when it is not cumulative to information already of record or being made of record in the application . . . .").

56. *See* Adenta GmbH v. OrthoArm, Inc., 501 F.3d 1364, 1374 (Fed. Cir. 2007) (agreeing with a district court's finding of no inequitable conduct based, in part, on the applicant's belief that the undisclosed information was cumulative).

57. *See* 37 C.F.R. § 1.56(c) (2008).

58. *Id.* § 1.56(c)(2).

59. *Id.* § 1.56(c)(3) ("Every other person who is substantively involved in the preparation or prosecution of the application and who is associated with the inventor, with the assignee or with anyone to whom there is an obligation to assign the application."). For example, a senior scientist who is not a listed inventor, but worked with the inventor on the invention's underlying chemistry, is under a duty to disclose. *See* Sython IP, Inc. v. Pfizer Inc., 472 F. Supp. 2d 760, 775 (E.D. Va. 2007) (identifying at least six individuals under a duty to disclose).

60. *See* FMC Corp. v. Hennessy Indus., Inc., 836 F.2d 521, 526 n.6 (Fed. Cir. 1987) ("As a general rule, there is no duty to conduct a prior art search, and thus there is no duty to disclose art of which an applicant could have been aware."). There have been recent cases that, arguably, establish a duty to inquire in limited circumstances as to possible prior art. *See* Brasseler, U.S.A. I, L.P. v. Stryker Sales Corp., 267 F.3d 1370, 1380

4. Remedy

The inequitable conduct doctrine is available as an affirmative defense to allegations of patent infringement.[61] An alleged infringer must prove, by clear and convincing evidence, that the applicant intentionally failed to disclose information material to the invention's patentability during patent prosecution.[62] If inequitable conduct is established, all the patent's claims are rendered unenforceable.[63] Depending on the circumstances, inequitable conduct with respect to a particular patent can infect and render unenforceable other related patents.[64] The doctrine has a much larger effect than a finding of invalidity for non-novelty or obviousness, which simply renders the particular claim in question invalid.[65]

C. Recent Critiques of the Doctrine

Since its inception,[66] the doctrine has garnered a tremendous amount of attention and criticism from the bar and the judiciary. The Federal Circuit has often noted the importance and seriousness of the doctrine.[67] Some Federal Circuit judges have gone out of their way to criticize the

---

(Fed. Cir. 2001) ("Where an applicant knows of information the materiality of which may so readily be determined, he or she cannot intentionally avoid learning of its materiality, even through gross negligence . . . ."). The court in *Brasseler* continued, however, noting that "[t]he mere possibility that material information may exist will not suffice to give rise to a duty to inquire . . . ." *Id.* at 1382.

61. *See* 35 U.S.C. § 282(1) (2006) (indicating that an alleged infringer can plead "unenforceability").

62. *See* GFI, Inc. v. Franklin Corp., 265 F.3d 1268, 1273–74 (Fed. Cir. 2001).

63. *See* Kingsdown Med. Consultants, Ltd. v. Hollister, Inc., 863 F.2d 867, 877 (Fed. Cir. 1988) (en banc).

64. *See* Consol. Aluminum Corp. v. Foseco Int'l Ltd., 910 F.2d 804, 812 (Fed. Cir. 1990).

65. *See* 35 U.S.C. § 282(2)–(3) (2006); Amazon.com, Inc. v. Barnesandnoble.com, Inc., 239 F.3d 1343, 1351 (Fed. Cir. 2001) (indicating that validity is determined on a claim-by-claim basis).

66. Many point to the Supreme Court's decision in *Precision Instrument Manufacturing Co. v. Automotive Maintenance Machinery Co.*, 324 U.S. 806 (1945), as the beginning of the inequitable conduct doctrine. Mack, *supra* note 20, at 152. The Supreme Court in *Precision Instrument* grounded inequitable conduct in the doctrine of unclean hands. *Precision Instrument*, 324 U.S. at 815–16.

67. *See Star Scientific*, 537 F.3d at 1366 ("[T]he severity of the penalty has not changed, and thus courts must be vigilant in not permitting the [inequitable conduct] defense to be applied too lightly.").

current use of the doctrine—viewing it as "overplayed"[68] and labeling its habitual assertion in litigation as "an absolute plague."[69]

The doctrine gets as much attention, if not more, from patent practitioners. Almost every patent Continuing Legal Education ("CLE") program includes a discussion of the doctrine.[70] The Practicing Law Institution ("PLI") issues multiple articles a year on recent developments in inequitable conduct law.[71] Even blog posts detailing recent inequitable conduct cases inevitably receive numerous comments from patent attorneys, postulating (and complaining) as to the breadth of the decision's impact.[72] This attention by practitioners is not surprising given that the doctrine focuses on "the person rather than the patent" by reviewing the patent attorney's actions to determine whether they engaged in inequitable conduct.[73]

This extensive attention from the courts and the bar has pushed the inequitable conduct doctrine into the general discussion about patent reform. Since early 2000, many commentators have focused on perceived shortcomings of the United States patent system. Highly publicized reports issued by the Federal Trade Commission ("FTC") in 2003 and the National Research Council ("NRC") in 2004 discussed target areas for reform.[74] Included in these discussions is the inequitable conduct doctrine. The in-

---

68. Kimberly-Clark Corp. v. Johnson & Johnson, 745 F.2d 1437, 1454 (Fed. Cir. 1984).

69. *See* Burlington Indus. v. Dayco Corp., 849 F.2d 1418, 1422 (Fed. Cir. 1988). This critique has continued in recent case law. *See* Aventis Pharma S.A. v. Amphastar Pharm., Inc., 525 F.3d 1334, 1349 (Fed. Cir. 2008) (Rader, J., dissenting) ("[I]nequitable conduct has taken on a new life as a litigation tactic.").

70. *See, e.g.*, American Intellectual Property Law Association, AIPLA 2008 Spring Meeting Schedule: Friday Schedule (May 16, 2008), *available at* http://www.aipla.org/Content/ContentGroups/Meetings_and_Events1/Spring_Meetings/20086/Program-Friday.pdf; University of Texas Continuing Legal Education, UTCLE 12th Annual Advanced Patent Law Institute, Program for Patent Law – Austin, http://www.utcle.org/ conference_overview.php?conferenceid=763 (last visited Feb. 4, 2009).

71. *See, e.g.*, Roxana H. Yang, *Duty of Disclosure & Inequitable Conduct—Who, What, When, & How?*, *in* ADVANCED PATENT PROSECUTION WORKSHOP 2007: CLAIM DRAFTING & AMENDMENT WRITING, at 557, 562 (PLI Patents, Copyrights, Trademarks, & Literary Prop., Course Handbook Series No. 11375, 2007).

72. *See, e.g.*, Posting of Dennis Crouch to Patently-O, What a Mess: Inequitable Conduct Based on Failure to Submit, http://www.patentlyo. com/patent/2007/05/ what_a_mess _ine.html (May 21, 2007 22:40 CST) (containing over a 110 comments to a post describing the holding in *McKesson Information Solutions, Inc. v. Bridge Medical, Inc.*, 487 F.3d 897 (Fed. Cir. 2007)).

73. Chisum, *supra* note 9, at 279.

74. FED. TRADE COMM'N, *supra* note 10; NAT'L RESEARCH COUNCIL, *supra* note 10.

equitable conduct doctrine has even garnered Congress's attention. Over the past four years, almost every draft of patent reform legislation contains some amendment to the doctrine.[75] Even the USPTO has proposed a new regulation concerning the applicant's duties regarding information submissions.[76]

From this discourse, two main critiques have emerged. One, supported by the NRC report, is that the doctrine is asserted too frequently by defendants and creates exorbitant litigation costs.[77] The second critique, discussed in the FTC report, focuses on the doctrine's failure to impose additional duties on the applicant, such as a duty to search or at least provide relevancy statements with regards to what is submitted.[78] Each noted problem has generated its own legislative solution, with the latter also prompting a proposed regulation by the USPTO. These two criticisms, the previously proposed legislation, and USPTO responses are discussed in detail below.

1. Creation of Unnecessary Litigation Costs and the Legislative Response

By recent estimates, the inequitable conduct defense is asserted in around one fourth of all patent cases filed.[79] More than one member of the

---

75. *See, e.g.*, Patent Reform Act of 2007, H.R. 1908, 110th Cong. (2007); Patent Reform Act of 2006, S. 3818 § 5, 109th Cong. (2006); Patent Reform Act of 2005, H.R. 2795 § 5, 109th Cong. (2005). The 2009 version of the Patent Reform Act does not modify the inequitable conduct doctrine, but there is still significant pressure to make statutory changes to the doctrine before the Act is adopted. *See supra* note 12.

76. *See* Changes To Information Disclosure Statement Requirements and Other Related Matters, 71 Fed. Reg. 38,808 (July 10, 2006) (to be codified at 37 C.F.R. pt. 1) (proposing to change the IDS requirements to include relevancy statements, in addition to other requirements).

77. *See* NAT'L RESEARCH COUNCIL, *supra* note 10, at 123.

78. *See* FED. TRADE COMM'N, *supra* note 10, at 12–13. The contents of a relevancy statement can vary, but usually includes some statement by the applicant as to how the submitted art is relevant to the patentability of one or more of the application's claims. *See, e.g.*, Mack, *supra* note 20, at 149.

79. *See Id.* at 156 ("Accused infringers, however, continue to plead the defense with regularity. Table 1 illustrates this regularity; from 2000 to 2004, an inequitable conduct adjudication appeared in 16% to 35% of all reported patent opinions."); Katherine Nolan-Stevaux, Note, *Inequitable Conduct Claims in the 21st Century: Combating the Plague*, 20 BERKELEY TECH. L.J. 147, 160–62 (2005) ("[I]t appears that parties frequently allege inequitable conduct where courts find no evidence of it.").

judiciary has viewed this rate as inappropriately high.[80] The NRC report and other commentators concur.[81]

This rate of pleading inequitable conduct is problematic because litigation of inequitable conduct claims is particularly costly. Most of the high cost comes from the subjective element of the doctrine—intent.[82] The circumstantial nature of most intent evidence makes summary judgment particularly difficult. Moreover, the deposition of the prosecuting attorney who handled the application is almost always necessary in the inequitable conduct inquiry.[83] Such depositions are uniquely costly because they are littered with complex attorney-client privilege issues that generate their own legal questions which demand additional attorney and judicial resources to resolve.[84]

Introducing inequitable conduct into the litigation also diverts attention from the heart of the dispute—the validity and infringement of the patent at issue. While inequitable conduct does concern the patent, the actual validity of the patent is irrelevant to the doctrine. Inequitable conduct inquiries turn into satellite litigations where the effort expended has little spillover benefits for other parts of the litigation. The time and energy spent on the defense may detract from the core issues and hamper their complete and correct resolution.[85] For these reasons, many push for reforms to lower the rate of pleading and reduce the cost of litigating inequitable conduct.

---

80. Kimberly-Clark Corp. v. Johnson & Johnson, 745 F.2d 1437, 1454; Burlington Indus., Inc. v. Dayco Corp., 849 F.2d 1418, 1422 (Fed. Cir. 1988).

81. *See* NAT'L RESEARCH COUNCIL, *supra* note 10, at 122 ("Another major complaint is that the defense is asserted too freely."); Nolan-Stevaux, *supra* note 79, at 148 ("The practice of asserting a defense of inequitable conduct, regardless of the merits of the defense in a given case, has reached the breaking point.").

82. See Doug Harvey, Comment, Reinventing the U.S. Patent System: A Discussion of Patent Reform Through an Analysis of the Proposed Patent Reform Act of 2005, 38 TEX. TECH. L. REV. 1133, 1152 (2006) ("Due to its subjective nature, the inequitable conduct defense is time consuming and expensive, and the abuse of the defense adds to the delays and increases the costs of litigation.").

83. Robert C. Faber, *Prosecution Ethics*, *in* ADVANCED PATENT PROSECUTION WORKSHOP 2008: CLAIM DRAFTING & AMENDMENT WRITING, at 13, 27 (PLI Patents, Copyrights, Trademarks, & Literary Prop., Course Handbook Series No. 14964, 2008) ("The time and cost of discovery will be increased by the need to investigate possible inequitable conduct and the associated discovery and, ultimately at trial, the cost of presenting the separate inequitable conduct defense.").

84. *See* Lynn C. Tyler, Kingsdown *Fifteen Years Later: What Does It Take to Prove Inequitable Conduct?*, 13 FED. CIR. B.J. 267, 269–70 (2003) (noting that, in extreme cases, inequitable conduct can cause the attorney-client privilege and work product immunity to be lost).

85. *See* Goldman, *supra* note 18, at 89.

The 2005 proposed changes to the inequitable conduct doctrine, contained in Section 5 of H.R. 2795, attempted to address the litigation cost concern.[86] The proposed legislation would have made the determination of inequitable conduct the exclusive province of the USPTO.[87] The doctrine would no longer be a defense to a claim of infringement. Instead, if inequitable conduct were alleged in litigation, the matter would be referred to the USPTO after the litigation ended.[88] The legislation also required a predicate finding that the non-disclosed information rendered one or more asserted patent claim invalid.[89]

The 2006 version of the patent reform legislation, S 3818, proposed modifications to the doctrine that were less dramatic than those in H.R. 2975. S. 3818 did require a predicate finding of invalidity,[90] but kept the doctrine in district court, not proposing the use of the USPTO to adjudicate such disputes.[91]

The changes in H.R. 2975, presumably, were intended to keep litigation costs down. Referral of the matter to the USPTO was one of the NRC report's proposals to "discourag[e] resort to the inequitable conduct defense and therefore reduc[e] its cost."[92] By making a finding of invalidity a prerequisite, costs would be reduced by limiting the instances under which the doctrine is litigated. This also necessarily heightens the materiality standard, making the doctrine tougher to plead. An alleged infringer must claim the undisclosed information renders one or more patent claims invalid, not that the information is simply relevant to the patentability issue.

2. *Lack of an Expanded Duty and the Legislative and USPTO Response*

Another criticism is the narrowness of the current duty to disclose. The critique is that the duty is improperly limited to providing only the infor-

---

86. *See* Patent Reform Act of 2005, H.R. 2795, 109th Cong. (2005).

87. Patent Reform Act of 2005, H.R. 2795, 109th Cong. § 136(a), (c) ("No court or Federal department or agency other than the [USPTO], and no other Federal or State governmental entity, may investigate or make a determination or an adjudication with respect to an alleged violation of the duty of candor and good faith . . . .").

88. *Id.* at § 136(c)(4).

89. That is, H.R. 2795 increases the standard of materiality to include only that information that actually results in a patent claim being held unpatentable. *Id.* at §136(d). Also, the examiner must have relied on the information in question when determining patentability. *Id.* at § 136(d)(3)(B).

90. Patent Reform Act of 2006, S. 3818, 109th Cong. § 5 (2006).

91. *See generally id.*

92. NAT'L RESEARCH COUNCIL, *supra* note 10, at 123.

mation already in the applicant's possession.[93] The applicant should do more. Suggestions range from requiring the applicant to search for additional material information and submit it to the USPTO to requiring the applicant to include "relevancy statements" indicating how the information she does submit is relevant to the patentability of the application.[94]

Critics believe that the current doctrine's narrow duty to disclose allows the applicant to simply bury his head in the sand.[95] In order to minimize the scope of his duty to disclose, the applicant affirmatively avoids coming across new information.[96] And to avoid misrepresenting information to the USPTO, the applicant says very little about what he does submit.[97] The USPTO is thus robbed of the additional knowledge a search would turn up and the insight that statements about the information would provide.

Two specific reasons are offered for requiring a search. First, if the applicant searches prior to filing her application, she may find out that her invention is not patentable.[98] Alternatively, she may discover that while patentable, the scope of protection she will get is narrow because the technological field is crowded.[99] If the invention sits in a field of art where

---

93. *See* Thomas Schneck, *The Duty to Search*, 87 J. PAT. & TRADEMARK OFF. SOC'Y 689, 704 (2005) (arguing that there should be a duty to search and submit the results to the USPTO).

94. Changes To Information Disclosure Statement Requirements and Other Related Matters, 71 Fed. Reg. 38,808 (July 10, 2006) (to be codified at 37 C.F.R. pt. 1) (setting forth the proposed modification to 37 C.F.R. § 1.98 (2008) to include "[a]dditional disclosure requirements" under § 1.98(a)(3)); FED. TRADE COMM'N, *supra* note 10, at 12–13 (recommending the inclusion of statements of relevance regarding submitted prior art).

95. Thomas, *supra* note 17, at 315 (noting that, because of potential liability under the inequitable conduct doctrine, "many applicants are discouraged from conducting prior art searches in the first place").

96. *Id.* ("Concerned that the failure to disclose a known reference will lead to the unenforceability of the patent, some applicants prefer to await the examiner's search results rather than consult the prior art themselves."); *see* Scott D. Anderson, Comment, *Inequitable Conduct: Persistent Problems and Recommended Resolutions*, 82 MARQ. L. REV. 845, 852–53 (1999).

97. *See* FED. TRADE COMM'N, *supra* note 10, at 12 (noting the fear that "slight errors in description could fuel claims of mischaracterization and inequitable conduct").

98. *See* Hal Gibson, Note, *In the Wake of* Enzo*: The Impact of the Federal Circuit's Decision on the U.S. Life Science Industry*, 41 SAN DIEGO L. REV. 903, 932 n.176 (2004) ("Many patent applicants (or more accurately, their attorneys) do conduct their own prior art search before filing their patent so as to better craft their own patent claims.").

99. *See* John M. Benassi, et. al., *Claim Construction and Proving Infringement: The Impact of* Phillips, Festo, *and Their Progeny*, *in* PATENT LITIGATION 2008, at 201 (PLI Patents, Copyrights, Trademarks, & Literary Prop., Course Handbook Series No. 14977,

there has already been a tremendous amount of technology patented, she will not be able to capture much in a new patent application. As a result of the information discovered in pre-filing search, the applicant may forgo patenting altogether and remove the burden of examining an invalid or worthless application. Or, the patent applicant will tailor the application to contain more realistic claims in light of the prior art, simplifying examination by the USPTO.[100]

Second, more information presented to the USPTO can improve the patent examination.[101] With more information describing the prior art, the USPTO has a better chance of correctly determining the patent's validity. This line of arguments is exhaustively explored below in Section III.C.[102] But, for now, it can simply be said that a duty to search is an information-producing mechanism that would result in a more through and accurate examination.

A duty to provide relevancy statements provides benefits similar to the last reason for a duty to search. Relevancy statements help the USPTO understand the submitted information in the context of patent application.[103] The USPTO does not need to spend as much time digesting the submitted information. Nor does the USPTO have to expend as much energy placing the submitted information in the context of the patentability of the claimed invention. Such statements facilitate a better and more efficient examination.

Recent legislative and USPTO proposals address this concern of an overly narrow duty to disclose. In the 2007 Patent Reform Bill, Congress gives the USPTO the ability to establish an applicant's duty to search and disclose the results of that search when filing a patent application.[104] The 2007 bill no longer focuses on the litigation cost reducing procedures in

---

2008) ("[T]he prior art search could be directed to ensure that the most important claim elements (and those most likely to be copied) are drafted as precisely as possible.").

100. *See* Changes To Practice for Continued Examination Filings, Patent Applications Containing Patentably Indistinct Claims, and Examination of Claims in Patent Applications, 72 Fed. Reg. 46,716, 46,720–21 (Aug. 21, 2007) (to be codified at 37 C.F.R. pt. 1) ("A number of patent applications contain a large number of claims, which makes efficient and effective examination of such applications problematic.").

101. *See* Schneck, *supra* note 93, at 694 (noting that, due to resource constraints, examiners miss relevant prior art, which leads to the issuance of invalid patents).

102. *See infra* Section III.C.

103. *See* Changes To Information Disclosure Statement Requirements and Other Related Matters, 71 Fed. Reg. 38,808, 38,810 (July 10, 2006) (to be codified at 37 C.F.R. pt. 1) (indicating that relevancy statements "are intended to provide meaningful information to the examiner"); FED. TRADE COMM'N, *supra* note 10, at 12 (noting testimony that relevancy statements leads to "better managed" and "quality enhanced" examination).

104. Reform Act of 2007, H.R. 1908, 110th Cong. § 5(a)–(b) (2007).

previous reform bills.[105] Instead, the focus is on expanding the duty governed by inequitable conduct.

The USPTO has proposed changes to the IDS requirements to require that applicants provide additional information, in certain circumstances, about the submitted information.[106] An applicant may be required to pinpoint representative portions of the submitted information, correlate these identified portions with the patent claims, and explain how the submitted information is different from other information already submitted.[107] The new rules would also require the applicant justify why the application is patentable in light of the submitted information depending on the number of submissions and the timing of the IDS's filing.[108]

D. Disconnects in the Current Discourse

Almost every facet of the patent system—Congress, the Federal Circuit, the USPTO, and patent practitioners—is concerned about the state of the inequitable conduct doctrine and lodge specific criticisms against the doctrine. The judiciary and NRC both assert that the doctrine is over-asserted and that this results in burdensome litigation costs. The FTC believes that the duty to disclose should be expanded.

This is the first disconnect in the current discourse—the two major critiques of the doctrine push in opposite directions. Concerns about litigation costs suggest weakening the doctrine, but increasing the duty to disclose means the doctrine should be strengthened. The conflict between these two critiques is difficult to resolve. A reduction of litigation costs necessarily means rejecting a duty to search. Giving the doctrine a broader reach makes it easier and more likely to be asserted as a defense to a claim of patent infringement.[109] This magnifies the harm the litigation cost critique is trying to minimize. From the other direction, the removal of the doctrine to the USPTO in order to reduce litigation costs would minimize any impact of a new duty to search. Moving the doctrine to a less favorable forum makes its assertion and enforcement less likely, weakening the substantive boost a new duty is meant to create.

---

105. Some of the proposed amendments arguably address litigation costs. *See, e.g.*, H.R. 1908 § 12(b), (c) (modifying the inequitable conduct doctrine by heightening the standard for intent and providing for less harsh remedies).

106. Changes To Information Disclosure Statement Requirements and Other Related Matters, 71 Fed. Reg. at 38,809.

107. *Id.* at 38,809–10.

108. *Id.*

109. *See* Cronin, *supra* note 19, at 1346 ("However, as the definition becomes more expansive there becomes more of an incentive for alleged infringers to charge inequitable conduct during litigation proceedings.").

The character of these two criticisms exemplifies another disconnect. Both of these concerns are utilitarian-focused, looking at how the doctrine impacts the patent system's goal of creating an optimal incentive to invent.[110] The traditional rationale for the inequitable conduct doctrine, in contrast, is the maintenance of ethical standards during prosecution.

The doctrine's equitable roots give the doctrine its moral bent. Since the mid-1800's, the judiciary has driven the development of the inequitable conduct doctrine as a creature of equity.[111] The Supreme Court identified the inequitable doctrine as the "equitable maxim that 'he who comes into equity must come with clean hands.'"[112] If a patent was born from fraud or deceit, then its holder cannot ask a court to enforce the patent.[113] The doctrine is seen as "a vehicle for affirmatively enforcing the requirements of conscience and good faith."[114]

In contrast, the alleged problems with the doctrine focus on the doctrine's impact on the optimal procurement and enforcement of patent rights. The litigation-costs argument is part of a larger movement to reduce the costs and uncertainty associated with patent litigation.[115] A reduction in such costs minimizes the likelihood a patent holder can improperly holdup a competitor practicing outside the area of valid patent protection.[116] The argument to broaden disclosure duties is focused on ensuring that only truly patentable inventions receive patent protection.[117] If patents are issued for inventions that are actually unpatentable, these "bad" patents will improperly deter competitors and follow-on innovators.[118] These problems are focused on the utilitarian goal of maintaining an optimal in-

---

110. *See, e.g.*, NAT'L RESEARCH COUNCIL, *supra* note 10, at 7 (concluding that one of the reforms needed, to created effective and efficient enforcement of patent rights, is the modification of subjective litigation elements, such as the inequitable conduct doctrine).

111. *See* Mack, *supra* note 20, at 152; Precision Instrument Mfg. Co. v. Auto. Maint. Mach. Co., 324 U.S. 806, 815 (1945).

112. *Precision Instrument*, 324 U.S. at 814.

113. *Id.*

114. *Id.*

115. *See* NAT'L RESEARCH COUNCIL, *supra* note 10, at 7, 123.

116. *See* Joshua D. Sarnoff, *Abolishing the Doctrine of Equivalents and Claiming the Future After* Festo, 19 BERKELEY TECH. L.J. 1157, 1200–01 (2004) ("Even when patents do not convey market power, patentees may exploit uncertainty regarding the scope of patents to deter competition by posing the threat of high-cost infringement litigation.").

117. *See* supra Section II.C.2.

118. *See* Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495, 1516 (2001) ("Certainly the issuance of bad patents has the potential to deter competition that should be lawful in some marginal cases."); Leslie, *supra* note 16, at 113–14. (arguing that "some invalid patents can deter market entry and decrease consumer welfare even without active enforcement").

centive to invent—providing protection where it is needed but not giving overprotection that does more harm to innovation than good.

A similar disconnect surfaces when comparing the current view of the inequitable conduct doctrine to the patent reform movement as a whole. One of the focuses of the movement is the optimal balance of patent protection and open competition. That is, providing patent protection where it is needed to prompt invention and innovation, but reigning in patent protection where such protection, on net, is detrimental to society.[119] The inequitable conduct doctrine, in contrast, focuses on the deontological ethics of the patent applicant's actions. The reform movement is results oriented, while the doctrine is focused on the means. This view of the doctrine finds no home in today's patent discourse.

III. FRAMING THE INEQUITABLE CONDUCT DOCTRINE AS A PATENT QUALITY MECHANISM

Scholars have written extensively on the inequitable conduct doctrine.[120] But the scholarship has not engaged in two basic, interrelated exercises that would greatly assist the discourse. First, no one has attempted a comprehensive, theoretical analysis of how the doctrine as a whole affects patent applicants, patent examination, and potential inventors. That is, they have not linked the doctrine to the general push to improve patent quality and reform the patent system. A majority of the scholarship, instead, is piecemeal—focusing on specific parts of the doctrine,[121] particular proposed statutory changes, and individual Federal Circuit cases.[122] Performing a fundamental analysis of the doctrine would provide a framework by which current and future reforms could be tested.

Second, and related, almost all the scholarship has kept the discussion tied to the doctrine's equitable roots, focusing on the doctrine as an ethical

---

119. *See, e.g.*, Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1614–30 (2003) (discussing how patents have positive or negative impact on different industries based on the nature of investment and discovery within each industry).

120. *See, e.g.,* Cedric A. D'Hue, *Disclosing an Improper Verb Tense: Are Scientists Knaves and Patent Attorneys Jackals Regarding the Effects of Inequitable Conduct?*, 14 U. BALT. INTELL. PROP. L.J. 121 (2006); Goldman, *supra* note 18; Robert A. Migliorini, *Lessons for Avoiding Inequitable Conduct and Prosecution Laches in Patent Prosecution and Litigation*, 46 IDEA 221 (2006).

121. *See, e.g.*, Cronin, *supra* note 19.

122. *See, e.g.*, *supra* note 19.

tool.[123] As just noted, these discussions are of little help when proposed changes need to be considered in the context of broader, utilitarian-justified reforms.[124] Thinking about the doctrine needs to be modernized and framed in the same terms as other targets of patent reform.[125]

This Part revisits the underlying rationale for the doctrine. Instead of focusing on ethics, this Part articulates the various ways in which the doctrine impacts the quality of the application and its examination.[126] The doctrine is well suited to affect quality of the process of issuing a patent given that the doctrine applies to all aspects of prosecution related to patentability. The beauty of placing inequitable conduct in the context of patent quality is that this analysis interjects the doctrine into the current discourse of patent reform. As demonstrated below, the doctrine can be an effective tool in improving patent quality, the system of patent examination, and the incentives generated by the system.

---

123. As previously discussed, a recent article has discussed the inequitable conduct doctrine's impact on patent quality but fails to fully develop the concept. *See* Mack, *supra* note 20, at 166–69.

124. One of the main thrusts of the patent reform movement is to ensure the proper balance between the incentive to invent and the ability to follow-on innovate is maintained. *See* FED. TRADE COMM'N, *supra* note 10, at Exec. Summ., 4–5. Part of this reform is to ensure that the only patents to issue from the USPTO are those that truly meet the patentability standards. *Id.* at 9–10.

125. *See, e.g.*, KSR Int'l Co. v. Teleflex, Inc., 550 U.S. 398, 427 (indicating that the nonobviousness requirement should not be set too low so as to allow patents to issue that "might stifle, rather than promote, the progress of useful arts"); FED. TRADE COMM'N, *supra* note 10, at 15 (noting competition concerns that should be considered when reforming the nonobviousness requirement).

126. Again, this is not to say that the inequitable conduct doctrine's impact on patent quality has never been mentioned. *See, e.g.*, Norton v. Curtiss, 433 F.2d 779, 794 (C.C.P.A. 1970) ("The highest standards of honesty and candor on the part of applicants in presenting such facts to the office are thus necessary elements in a working patent system. We would go so far as to say they are essential."); 37 C.F.R. § 1.56(a) (2008) ("The public interest is best served, and the most effective patent examination occurs when, at the time an application is being examined, the Office is aware of and evaluates the teachings of all information material to patentability."); Mack, *supra* note 19, at 166–69; Rene D. Tegtmeyer, *The Patent and Trademark Office View of Inequitable Conduct or Attempted Fraud in the Patent and Trademark Office*, 16 AIPLA Q.J. 88, 88 (1988) (quoting former Assistant Commissioner of the USPTO that "[t]he purpose of the duty of disclosure requirement, as the Patent and Trademark Office (PTO) views it, is to improve the quality of examination and the validity of patents by assuring that material information is called to the examiner's attention and considered in the patent examining process"); Thomas, *supra* note 17, at 313–14 (labeling the inequitable conduct doctrine as an "information-gathering technique[]").

Again, none of these articles has provided a detailed theory as to exactly how the doctrine can impact patent quality.

In order to frame the doctrine as a patent quality tool, this Part first defines patent quality and links the quality concerns to the need for information. This Part also discusses the lack of an inherent incentive for applicants and others to provide the USPTO with relevant information during examination. This Part then explains how the inequitable conduct doctrine provides such an incentive. The doctrine improves the quality of the application by increasing the patent attorney knowledge and understanding of the invention and the related technological area. The doctrine also helps ensure that the application and related correspondence are drafted with care. The doctrine operates as an information producer and verifier, giving the USPTO more resources and time to properly examine the application.

A. Patent Quality and Information

Put simply, optimal patent quality is the issuance of patents that meet the patent requirements and the rejection of those that do not. The assurance of a good patent quality is all about information—both access to it and time for the examiners to use it.

1. Patent Quality Problem Defined

The concept of patent quality focuses on the patentability of those patent claims allowed by the USPTO. The patent system assumes that only those patent applications that describe and claim a patentable advance are granted the power to exclude. Those patents that meet the validity requirements—claim useful, novel, and nonobvious inventions that are fully disclosed and enabled—are considered to be of good quality.[127] Patents issued by the USPTO that claim subject matter that does not meet the patentability requirements are known as poor quality, or "bad" patents.[128]

Most agree that the patent system should maintain high patent quality.[129] Granting patents on patentable advances provides incentives for the

127. *See, e.g.*, Lichtman & Lemley, *supra* note 15, at 46–49.

128. *See* Jay P. Kesan & Andres A. Gallo, *Why "Bad" Patents Survive in the Market and How Should We Change? — The Private and Social Costs of Patents*, 55 EMORY L.J. 61, 63 (2006) ("The common criticism from all sides is that the Patent Office grants patent claims that are broader than what is merited by the invention and the prior art, resulting in so-called "bad" or improvidently granted patents."); Lee Petherbridge, *Positive Examination*, 46 IDEA 173, 175 (2006) ("Questionable, or low quality, patents are those patents that should never have issued from the Patent Office because they fail to meet the statutory requirements for patentability.").

129. There is debate, however, as to how much resources should be allocated to ensuring that valid patents are granted by the USPTO. *Compare* Lemley, *supra* note 130, at 1497 (arguing that few resources should be expended in improving examination "[b]ecause so few patents are ever asserted against a competitor, it is much cheaper for society to make detailed validity determinations in those few cases than to invest addi-

creation of beneficial technical advances and facilitates their commercialization.[130] Society benefits when quality patents issue.

In contrast, the issuance of patents of poor quality has deleterious effects. A bad patent, for example, may give its holder exclusive control over a minor technological advance, creating roadblocks to innovation typically allowed under patent law.[131] Since even poor quality patents enjoy a presumption of validity, holders of bad patents have the power to impede the legitimate innovation of others and seek licensing fees for activities that are actually allowable.[132] The bad patent creates *in terrorem* effects, deterring socially acceptable and beneficial behavior.[133] Those who want to use the patented technology must expend significant resources to determine and, if forced, legally establish, that the patent is invalid.

Many factors contribute to the poor quality of U.S. patents. Some critiques point to the standards for determining patentability, concluding that they are too low and, even if properly applied, result in the issuance of socially detrimental patents.[134] Most, however, view the patent quality problem as an information and resource problem. That is, the USPTO does not

---

tional resources examining patents that will never be heard from again"), *with* Shubha Ghosh & Jay Kesan, *What Do Patents Purchase? In Search of Optimal Ignorance in the Patent Office*, 40 HOUS. L. REV. 1219, 1225–26 (2004) (arguing that changes in the examination process can be a cost-beneficial way of improving patent quality).

130. Mark A. Lemley, *Ex Ante versus Ex Post Justifications for Intellectual Property*, 71 U. CHI. L. REV. 129, 129-30 (2004) (describing patent law's ability to create an ex ante incentive to invent); Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 276–78 (1977) (describing an ex post theory of patents where protection assist the development of the patented invention).

131. *See* 3 ROBERT MERGES & JOHN DUFFY, PATENT LAW AND POLICY: CASES AND MATERIALS 647 (3d ed. 2002); Christopher A. Cotropia, *Patent Law Viewed Through an Evidentiary Lens: The "Suggestion Test" as a Rule of Evidence*, 2006 B.Y.U. L. REV. 1517, 1525 ("Exclusive control over these minor developments would act as roadblocks, creating disincentives to future inventors. Many patents on small technical advances make it extremely difficult and 'expensive to search and to license' these patents in order to produce further innovations.").

132. *See* Lichtman & Lemley, *supra* note 15, at 47–48 (noting that the presumption of validity makes "defendants face an uphill battle persuading the courts to overrule that errant determination").

133. *See* John R. Thomas, *The Responsibility of the Rulemaker: Comparative Approaches to Patent Administrative Reform*, 17 BERKELEY TECH L.J. 727, 731 (2002) (detailing these detrimental effects).

134. The debate surrounding the nonobviousness requirement provides a good example of this type of discussion. *See, e.g.*, John F. Duffy, *Inventing Invention: A Case Study of Legal Innovation*, 86 TEX. L. REV. 1 (2007) (discussing the various standards for determining nonobviousness and how they implicate differing policy views on patent law).

have access to adequate information to correctly determine whether a claimed invention is novel and nonobvious.[135] This is particularly problematic in new technological areas, such as software and business methods, where the best information on what has previously been done is not in prior patents, but trade publications, public presentations, product brochures, and computer code.[136]

Even if the USPTO does have access to such information, patent examiners often do not have the time to find and apply it to the patent claims.[137] The number of patent applications is rising exponentially each year while, at the same time, the USPTO faces a significant examiner attrition rate.[138] Examiners are given very little time to perform a complete examination—gain an understanding of the invention, determine the meaning of the patent claims, search the prior art, apply the prior art to the claims, write office actions, and respond to the applicant's arguments multiple times.[139]

Finally, some worry that poorly drafted patents hinder the examiner's ability to understand the claimed subject matter.[140] The harder the applica-

---

135. *See* Lichtman & Lemley, *supra* note 15, at 46 ("Information is a second significant impediment to PTO review.").

136. *See* Thomas, *supra* note 17, at 318–19. Thomas describes the unique information problem as it relates to new technologies:

> For software, business methods, and other postindustrial inventions, the repository of issued patents insufficiently samples the prior art. Examiners who primarily rely upon the patent literature to generate prior art in these fields are quite likely to allow patents to issue based upon information already within the public domain. Even those diligent examiners who consult the nonpatent literature might be limited to a sparse prior art collection.

Id.

137. *See* Lichtman & Lemley, *supra* note 15, at 46 (identifying the resource problem faced by the USPTO to effectively review the growing number of applications); Thomas, *supra* note 17, at 314 ("[T]he average time allocated for an examiner to address one application is understood to be between sixteen and seventeen hours. Given the complexities involved in parsing an application, conducting a prior art search and drafting an Office Action, this period is surprisingly short.").

138. *See* Beth Simone Noveck, *"Peer to Patent": Collective Intelligence, Open Review, and Patent Reform*, 20 HARV. J.L. & TECH. 123, 132 (2006) ("[T]he USPTO still cannot hire quickly enough to keep pace with both the demands of the job and the attrition rate.").

139. Thomas, *supra* note 17, at 314 (noting that examiners are allotted between sixteen to seventeen hours per application).

140. Petherbridge, *supra* note 128, at 181–83, 192 ("[T]he better the Patent Office collects and uses information about the boundaries of the property right, the higher the quality of examination.").

tion is to comprehend, the more difficult it is for the examiner to properly and efficiently examine the application.[141] As the saying goes—garbage in, garbage out.

2. *Disincentives for Those Outside the USPTO to Solve the Quality Problem*

Why is the patent quality problem not self-correcting? Surely patent applicants have an interest in high patent quality. If the USPTO is doing a good job examining patents, a patent holder can readily rely on the USPTO's determination and not expend resources in making its own assessment after issuance. In turn, a patent holder can charge more for a clearly valid patent and valid patents are less likely to get embroiled in costly litigation and, thus, more efficient to enforce. Put simply, the value of a quality patent is higher than a bad one. Why wouldn't patentees want that?

Patent applicants have countervailing strategic reasons to ignore patent quality. As mentioned, even bad patents provide value to the holder because of the costs they create for others.[142] Any attempt to assist in improving patent quality may destroy a bad patent's value altogether by preventing it from ever issuing.[143] An issued, poor quality patent is more valuable than no patent at all.

Even for good patents, it may be in applicants' best interest to keep patent quality information to themselves. Applicants are in the best position to determine the true validity of the patent because of the information asymmetry between the inventor and potential challengers.[144] The inventor and related individuals know the most about the invention and potential prior art in the invention's technical field.[145] Thus, the patentee can make its own determination as to the quality of the patent. While costly, this de-

---

141. *Id.*

142. *See* Leslie, *supra* note 16, at 113–28 (detailing the many ways invalid patents "injure competition").

143. *See* R. Polk Wagner, *Reconsidering Estoppel: Patent Administration and the Failure of* Festo, 151 U. PA. L. REV. 159, 215 (2002) (arguing that the patentee is incentivized to not provide prior art to the USPTO to "increase[e] the possibility that the PTO will 'miss something' and allow the unwarranted scope").

144. *Id.* at 214 ("Given the asymmetry of information, the incentives for a patentee to fail to produce relevant information are substantial.").

145. *See* Cotropia, *supra* note 22, at 84 ("The information in the specification is produced by the inventor, the lowest cost source for invention-specific information."); Wagner, *supra* note 143, at 213 ("Among the 'parties' to the patent transaction, the patentee is either the best informed or the one who can most easily and cheaply become the best informed about the context of her innovation.").

termination is more difficult, if not impossible, for those without easy access to information the patentee holds.[146] This information asymmetry gives the patentee the ability to engage in strategic behavior by withholding information and preventing a potential licensee or defendant from knowing the true value of the patent.

Irrespective of information asymmetry and strategic behavior, the cost of improving the quality of examination alone may deter applicants from engaging in self-help. Assisting in the examination process by either doing a pre-filing search for prior art or submitting prior art to the USPTO is a costly endeavor.[147] Some applicants are willing to take the risk of receiving a bad patent given the high costs of ensuring the patent is a good one.

Finally, other applicants may be ignorant of the patent quality situation at the USPTO.[148] An examination for patentability is what an applicant pays for and some may assume that is what they will get.

Third parties have no incentive to assist in the examination process. Mechanisms do exist for third parties to participate in an ongoing examination or to force the reexamination of an issued patent.[149] However, because challenges to validity "exhibit the characteristics of public goods," potential challengers face a collective action problem.[150] A successful challenger cannot prevent others from free-riding on the resulting patent invalidation by practicing the previously exclusive invention.[151] A potential challenger is better off keeping the invalidity information to herself and only using it when she is accused directly of infringement.[152]

Because of these reasons, the patent system cannot rely on applicants or third parties to *sua sponte* assist the patent examination process. Information-forcing rules must be considered.[153] That is, certain patent doc-

146. Wagner, *supra* note 143, at 215 (indicating that an applicant will not produce "the sort of information that might allow the PTO and the public to more usefully evaluate the scope of the patent").

147. Lemley, *supra* note 130, at 1510 (reporting the average cost of a prior art search as between $5000 and $7000).

148. *See, e.g.*, Lough v. Brunswick Corp., 86 F.3d 1113, 1122 (Fed. Cir. 1996) (recognizing that some inventors do not have a sophisticated understanding of the patent law system).

149. 35 U.S.C. § 135 (2006) (interference proceedings); 35 U.S.C. §§ 301–307 (2006) (ex parte reexamination); 35 U.S.C. §§ 311–318 (2006) (inter partes reexamination); *see also* Thomas, *supra* note 17, at 326–28 (detailing these various avenues of third party challenges, including a proposed post-grant opposition system).

150. Thomas, *supra* note 17, at 333.

151. *Id.* at 333–34.

152. *Id.* at 334.

153. *See* Wagner, *supra* note 143, at 216–17, 221 (discussing penalties for underproduction of information and viewing the prosecution history estoppel doctrine as one of

trines should force the patent applicant to act against their strategic interest. The inequitable conduct doctrine serves in this capacity by incentivizing applicants to produce valuable information and, in turn, improve patent quality.

B. Doctrine's Ability to Improve the Quality of Information before USPTO

The inequitable conduct doctrine is a disclosure doctrine, which, by its inherent nature, creates a flow of information from the applicant to the USPTO. The doctrine does even more by focusing on the production of information relevant to patentability and, in turn, *de facto* verifying it. Both of these aspects of the doctrine improve patent quality and are discussed below.

1. Produces Relevant Information to the USPTO

At its core, the doctrine is an information producer. The inequitable conduct doctrine requires patent applicants provide the USPTO with information relevant to the patentability of the claimed invention.[154] The doctrine acts as a conduit through which information from the patent attorney, the inventor, and related parties flows to the examiner. This information, coming in the form of patents, periodicals, data, physical specimens, affidavits, and the like, is directly related to patent examiner's primary responsibility—determining the application's patentability.

This information is especially relevant because the doctrine draws it from those who know the most about the invention and its area of technology—the inventor and those directly involved in the patent's prosecution.[155] These are all individuals—the bench scientists, technicians, technology group leaders, in-house patent attorneys, etc.—who were either intimately involved in the invention's creation or in the drafting of the patent application.

Getting information from these individuals gives the examiner access to information that is not contained in the databases readily available to her. Patent examiners have the ability to search world-wide patent databases and some technical article databases. However, they do not have

---

these penalties); *cf.* Scott R. Boalick, *Patent Quality and the Dedication Rule*, 11 J. INTELL. PROP. L. 215, 221 (2004) (arguing that a strong dedication doctrine "improve[s] overall patent quality by creating incentives for good patent drafting at the earliest stages of the patent acquisition process, and long before litigation arises").

154. *See supra* Section II.A.

155. *See supra* note 59 (identifying the various individuals beyond just the inventor who have a duty to disclose under the inequitable conduct doctrine).

ready access to all technical literature, such as specialized industry publications or dissertations, or the technologies themselves, such as computer code listings or actual devices.[156] The USPTO does not have access to information in the technology areas new to patenting.[157] Patent examiners also must rely on applicants to inform them of potential offers to sale, conference presentations, test data, and product brochures regarding the invention.[158]

The doctrine generates valuable information by placing information production responsibilities on a low-cost provider. Production of information costs the applicant,[159] but the doctrine limits this cost by requiring the applicant to consider only the information already in her possession.[160] More importantly, the cost to the applicant is lower than the cost of the examiner finding the same information. The examiner has some specialized knowledge, but less than the applicant in understanding the particular invention and processing the information already in the applicant's possession.[161] Even if the examiner can gain access to information similar to that which is at the disposal of the applicant, the examiner starts from scratch in evaluating the information's relevance to patentability. For at

156. Empirically, examiners are at a "disadvantage in searching for non-patent prior art and foreign patents." Bhaven N. Sampat, Determinants of Patent Quality: An Empirical Analysis 3 (Sept. 2005) (unpublished manuscript), *available at* http://siepr.stanford.edu/ programs/SST_Seminars/patentquality_ new.pdf_1.pdf. Notably, the information examiners have access to is growing everyday. *See* MPEP § 901.06(a) (2007) (describing the resources available to patent examiners); The Scientific and Technical Information Center ("STIC"), available at http://www1.uspto.gov/web/ offices/pac/dapp/sir/ stic/ brochure. html (same); Iain M. Cockburn et al., *Are All Patent Examiners Equal? The Impact of Characteristics on Patent Statistics and Litigation Outcomes*, (Nat'l Bureau of Econ. Research, Working Paper No. 8980, 2002), *available at* http://www.nber.org/ papers/w8980 (describing the STIC). A vast amount of inventor specific material, or hard to find material in a given field, is just not accessible.

157. Peter S. Menell, *A Method For Reforming the Patent System*, 13 MICH. TELECOMM. & TECH. L. REV. 487, 504 (2007) ("Similarly, since this is a new patent field, examiners have relatively little training in this area, there is little or no patent prior art, and time and database constraints severely limit the ability of examiners to search non-patent prior art."); Cockburn, *supra* note 156, at 6 ("In very young technologies, or in areas where the USPTO has just begun to grant patents, there may be very limited patent prior art.").

158. This information is unlikely to be found in any database available at the USPTO. *See, e.g.*, *supra* note 156.

159. *See infra* Section IV.C (detailing the costs of submitting information to the USPTO).

160. *See supra* Section II.A (noting how the current doctrine does not include a duty to search).

161. *See supra* note 145 (explaining how the inventor is in the best position to know the most about her invention).

least some information, the applicant has already filtered out the irrelevant material to force the examiner to repeat the process would be wasteful.

The breadth of relevant information under the doctrine—information that simply creates a prima facie case of invalidity but does not necessarily render the claim invalid—has a second-order information production effect. The exact information produced by an applicant may not be used by the examiner, but that information can mark a path to a different technological area that contains relevant information.[162] The information may also contain or prompt a line of technical thought that could justify a rejection.[163]

By mandating production of this valuable information, the inequitable conduct doctrine addresses some of the causes of the current patent quality problem. The doctrine provides the examiner with more invention-specific information from which sources examiners likely do not have access. The better the information, the better the examination.[164] Furthermore, the burden of producing this information is not born by the examiner, so it increases the amount of time the examiner has to complete the examination.

2. *Verifies Information Provided to the USPTO*

The doctrine also works as an information verifier. The inequitable conduct doctrine fundamentally requires that patent applicants must be truthful in their correspondence with the USPTO.[165] They cannot misrepresent information. Nor can they omit anything relevant to the truthfulness of disclosed information. Because of this duty of truthfulness and full disclosure, an examiner does not have to question the veracity of a statement or response by an applicant. They are self-authenticating.

---

162. *See, e.g.*, *In re* ICON Health & Fitness, Inc., 496 F.3d 1374, 1379–80 (Fed. Cir. 2007) (looking to folding bed art to invalidate an application directed toward a folding treadmill).

163. For example, examiners reject a claim as obvious if they find a reason to combine the prior art. *See* KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 416–17 (2007).

164. *See* Michael Astorino, *Obviously Troublesome: How High Should the Standard Be for Obtaining a Patent*, 89 J. PAT. & TRADEMARK OFF. SOC'Y 239, 250 (2007) ("The better the search the better the prior art rejections."); Nolan-Stevaux, *supra* note 79, at 159–60 ("[I]nequitable conduct [claims in patent law] also function[] as a penalty default to discourage applicants from playing strategic games."); Posting of Dennis Crouch to Patently-O, Evidence Based Prosecution: Non-Patent Art Leads to Rejections, http://www.patentlyo.com/patent/2006/10/evidence_based__3.html (Oct. 29, 2006 22:03 CST) (finding that most rejections are based on non-patent art).

165. *See* Li Second Family Ltd. P'ship v. Toshiba Corp., 231 F.3d 1373, 1378 (Fed. Cir. 2000) ("[A]ffirmative misrepresentations of material facts, failure to disclose material information, or submission of false material information, coupled with an intent to deceive, constitutes inequitable conduct.").

The doctrine's information verification function comes at a low cost. The individuals who make statements—the applicant, patent attorney, and related parties—are the ones who have to stand by their reliability. And since the applicant makes the statements, she is in the best position to attest to their accuracy. In contrast, the patent examiner is in a very poor position to determine veracity. Almost all correspondence is done in writing, removing the option for examiners to look for visual or audible signs of a particular statement's truthfulness.[166] With their heavy workload, examiners do not have the time to independently verify all of an applicant's statements and claims.[167] Nor do examiners have the resources or the training.[168] Verification costs are further minimized by requiring the applicant to attest only for information in her possession.[169] No duty to search for relevant art or other information exists currently, and thus, the applicant is not asked to expend the time and resources to attest to all the information or knowledge in a given area.

The doctrine also works as an external verifier of information. The threat of unenforceability not only assures the examiner that statements made by the applicant are true, it also assures others external to the USPTO that the information is correct.[170] Members of the public who may be looking at the patent as either an educational tool or an indicator of the patent holder's technological direction can rely on the inequitable conduct doctrine to settle any question as to the application's truthfulness. The applicant, as was the case when compared to the examiner, is in a much lower cost position to verify this information than each member of the public.

Problematically, the truth of any given submission is only as good as the subjective belief of the submitter.[171] The intent requirement of the inequitable conduct doctrine requires that the applicant only subjectively

166. Examiners do, on occasion, correspond with applicants by phone and in interviews. *See* 37 C.F.R. § 1.133 (2008) (establishing rules regarding examiner interviews).

167. They can make a request for additional information. 37 C.F.R. § 1.105 (2008). But, these are rarely used, most probably because making one uses up scarce examination time.

168. *See, e.g.*, U.S. Patent and Trademark Office, USPTO Patent Training Academy (PTA) (Mar. 6, 2006), http://usptocareers.gov/pdf/PatentTraining Academy1.pdf (listing the areas in which examiners are trained).

169. *See supra* Section II.A.

170. *See* Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625, 658 (2002) ("At the very least, investors can be assured that firms will not make objectively false statements in the body of the patent; if they do, they will bear both actual and reputational costs.").

171. *See supra* Section II.B.3.

believe that what is submitted is true.[172] It could turn out that the applicant was, objectively, wrong in her conclusion—the information may actually be incorrect. So, the doctrine's ability to verify information is only as good as the applicant's subjective knowledge. If the information is actually false, a reliant examiner could be lead astray by the erroneous comment. However, while this is certainly possible, the applicant is in the best position to have a correct understanding of how her invention works when she came up with the invention, the result of any tests done on the invention, and so on. Other statements, such as those in affidavits, are not meant to prove anything more than the affirmant's subjective belief. The applicant's subjective believe is, in most instances, the best the system can produce and the inequitable conduct doctrine ensures that is what is communicated.

This verification function improves patent quality. The quality of the information before the examiner is increased because it is much more likely to be true under the doctrine. Moreover, the examiner does not have to waste any of the finite examination time on making truth determinations.

C. Doctrine's Ability to Improve the Quality of the Patent Application

The doctrine also improves the quality the patent application and other correspondence with the USPTO. It does this by increasing the patent attorney's knowledge of the invention and related technology and causing the attorney to exercise more care when drafting the application and correspondence. By improving the quality of these documents—both their technical fidelity and accuracy—the doctrine can improve the quality of examination and the issued patent itself.

1. Increases the Patent Attorney's Knowledge of the Invention and Related Technology

The patent attorney, in the process of complying with the inequitable conduct doctrine, gains a great deal of knowledge about the invention and its technological field. Initially, compliance generates a base of knowledge in the relevant technology. To assess materiality, the patent attorney must read all of the information within her possession to determine its relevance

172. *See* Liquid Dynamics Corp. v. Vaughan Co., 449 F.3d 1209, 1227 (Fed. Cir. 2006) ("Intent [to deceive the USPTO] is a subjective inquiry into whether the inventor knew the information was material and chose not to disclose it.").

to patentability.[173] The doctrine's focus on patentability information forces the attorney to concentrate on information related to the invention. Take the patent at issue in *Monsanto Co. v. Bayer Bioscience N.V.*, which claimed a specific type of genetically-modified corn that is toxic to insects, but not humans.[174] The inequitable conduct doctrine would require the patent attorney who prosecuted the patent to read and analyze the information within her possession discussing the genetic modification of food and safe pesticide products in order to determine what needs to be disclosed. Through the process, she would learn the chemistry and biology behind genetic modification and pesticides. She would also learn the composition of previous pesticides, how they were designed, and their particular uses.[175] The patent attorney digests this type of information not only to gain an understanding of the technological area, but also to comply with a legal doctrine. This added importance means that not only will patent attorneys read the information, but they will do so with care and attention to detail.

In complying with the doctrine, the patent attorney also learns more about the invention itself. The patent attorney must obviously speak to the inventor in order to draft the application and get an understanding of what she can claim as the invention.[176] The inequitable conduct doctrine, however, forces her to dig deeper and analyze all information regarding the invention. She must evaluate all publications, correspondence, and prior uses regarding the invention to see if there is any material information—such as a public use or an offer for sale—she must disclose.[177] The patent attorney must also ensure that all of the statements in the application re-

---

173. The inequitable conduct doctrine inquires as to whether a piece of information meets a threshold level of materiality. *See* Dayco Prods., Inc. v. Total Containment, Inc., 329 F.3d 1358, 1362–63 (Fed. Cir. 2003).

174. *See* Monsanto Co. v. Bayer Bioscience N.V., 514 F.3d 1229, 1237–-39 (Fed. Cir. 2008).

175. *See id.* at 1238–39.

176. An application is not complete until the inventor signs an oath declaring, amongst other things, that "[s]tates that the person making the oath or declaration acknowledges the duty to disclose to the Office all information known to the person to be material to patentability as defined in § 1.56." 37 C.F.R. § 1.63(b)(3) (2008). The USPTO recently stated that it would reject oaths that "do not expressly acknowledge a duty to disclose information material to patentability." U.S. Patent & Trademark Office, Duty of Disclosure Language Set Forth in Oaths or Declarations Filed in Nonprovisional Patent Applications (Jan. 2, 2008), *available at* http://www.uspto.gov/web/offices/pac/dapp/opla/ preognotice/duty_of_disclosure.pdf.

177. This information would be relevant to patentability in light of the on-sale bar. *See* 35 U.S.C. § 102(b) (2006); Brasseler, U.S.A. I, L.P. v. Stryker Sales Corp., 267 F.3d 1370, 1376–77 (Fed. Cir. 2001) (discussing the materiality of on-sale information).

garding the invention's operation, prior testing, and the invention's construction are correct.[178]

The doctrine is also structured to funnel information from the inventor and related parties to the patent attorney. The duty to disclose is not imposed only on the patent attorney communicating with the USPTO. The inventor and "[e]very other person who is substantively involved in the preparation or prosecution of the application and who is associated with the inventor" must also disclose material information.[179] Material information in the hands of any such individual necessarily makes its way to the patent attorney before it is disclosed to the USPTO.[180] The patent attorney is the hub for all communications to and from the USPTO.[181] The doctrine's broad scope thus causes those most knowledgeable of the invention and relevant technology to share their knowledge with the patent attorney to meet their duty to disclose.[182] As a result, the drafter of the application is exposed to even more relevant information.

All this additional knowledge translates into a patent application that is easier to examine. Many patent law doctrines ask the USPTO and courts to view the patent through the lens of one skilled in the relevant technological art.[183] A knowledgeable patent attorney can write to this intended audience because she has gained an understanding of a given science or area of engineering and knows the relevant terminology.[184] This means the

---

178. This information would be relevant to the disclosure requirements. *See* 35 U.S.C. § 112 ¶ 1 (2006).

179. 37 C.F.R. § 1.56(c) (2008).

180. This is typically done by an over-inclusive request by the patent attorney asking the relevant parties if they know of any information related to the invention's subject matter.

181. *See* 37 C.F.R. §§ 1.33, 1.34 (2008) (noting that patent attorneys act as a representative of the inventor filing for an application).

182. *See supra* note 147 (describing how the inventor has the greatest knowledge of the invention).

183. Patent claims—which define the scope of exclusivity—are interpreted as the terms are understood by one skilled in the art. *See* Phillips v. AWH Corp., 415 F.3d 1303, 1313 (Fed. Cir. 2005) (en banc) ("[T]he ordinary and customary meaning of a claim term is the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention, i.e., as of the effective filing date of the patent application."). "The enablement requirement is satisfied when one skilled in the art, after reading the specification, could practice the claimed invention without undue experimentation." AK Steel Corp. v. Sollac, 344 F.3d 1234, 1244 (Fed. Cir. 2003) (citing *In re* Wands, 858 F.2d 731, 736–37 (Fed. Cir. 1988)).

184. *See* Patricia Wright, *Writing Technical Information*, 14 REV. RES. EDUC. 327, 339–40 (1987) (discussing the knowledge needed to be an effective technical writer).

examiner, who specializes in the invention's technological field,[185] spends less time trying to understand the application and what is claimed and more time determining patentability.

The patent attorney's better understanding of what has been previously accomplished in a technical field also facilitates the drafting of patent claims that avoid subject matter that is not novel or obvious.[186] This streamlines examination because unpatentable subject matter is weeded out prior to filing. The examiner does not waste time rejecting clearly invalid claims.

The patent application also becomes more socially beneficial. A patent is not only meant to incentivize the creation of the invention; it is also meant to educate others and facilitate improvements and design-arounds of the claimed invention.[187] The doctrine causes the patent drafter to better understand the invention and be able to "talk the talk" technically. And the more technically-accurate and accessible the patent, the better it can fulfill these goals. The patent becomes like any other scientific reference material, explaining the subject matter in a way that is comprehensible to its intended audience.[188]

---

185. *See* Am. Hoist & Derrick Co. v. Sowa & Sons, Inc., 725 F.2d 1350, 1359 (Fed. Cir. 1984) (indicating that examiners "are assumed to have some expertise in interpreting the references and to be familiar from their work with the level of skill in the art").

186. Patent attorneys routinely try to anticipate rejections by the USPTO when prosecuting patent applications in order to avoid them. *See, e.g.*, Rajiv P. Patel et. al., *Understanding After Final and After Allowance Patent Practice*, *in* ADVANCED PATENT PROSECUTION WORKSHOP 2008: CLAIM DRAFTING & AMENDMENT WRITING, at 805, 836 (PLI Patents, Copyrights, Trademarks, & Literary Prop., Course Handbook Series No. 14964, 2008) ("With careful planning and analysis, an applicant can anticipate and address these rejections in one or more different manners -- and in some instances avoid such rejections altogether.").

187. *See* Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 152 (1989) ("[A]fter the expiration of a federal patent, the subject matter of the patent passes to the free use of the public as a matter of federal law."); *see also* United States v. Dubilier Condenser Corp., 289 U.S. 178, 186 (1933) ("An inventor deprives the public of nothing which it enjoyed before his discovery, but gives something of value to the community by adding to the sum of human knowledge."); State Indus., Inc. v. A.O. Smith Corp., 751 F.2d 1226, 1236 (Fed. Cir. 1985) ("One of the benefits of a patent system is its so-called 'negative incentive' to 'design around' a competitor's products, even when they are patented, thus bringing a steady flow of innovations to the marketplace.").

188. *See* Dorothy A. Winsor, *Engineering Writing/Writing Engineering*, 41 COLLEGE COMPOSITION & COMM. 58, 58 (1990) ("We talk, therefore, of language, and particularly written language, as a tool for constructing ideas, of a given field of knowledge being created by the interaction of its practitioners' texts, and of knowledge itself, including scientific knowledge, as rhetorically shaped.").

Gaining a better understanding of the invention also allows the patent attorney to draft claims that give the inventor and her company the necessary "shelf space" so they can effectively commercialize the invention.[189] Patent protection incentivizes invention because it gives the inventor an ability to recoup her research and development costs.[190] Patents do this by giving the patent holder the ability to exclude competitors and control price.[191] The less understanding the patent attorney has of the invention, the less likely she will draft claims that facilitate this purpose of patents and, in turn, are valuable to her client. Patenting is as much a business decision as it is a legal one. By giving the patent attorney more information about the invention and its use prior to filing, the doctrine allows the patent application to better align with the invention's intended commercial use.

2. *Increases the Care Taken in Drafting the Application and Related Correspondence*

The doctrine also prompts the patent attorney to exercise more care in drafting the patent application and correspondence with the USPTO. The inequitable conduct doctrine penalizes applicants for misleading the USPTO with false statements in patent applications or other correspondence.[192] The choice of a single word can make the difference between full

---

189. *See* Kitch, *supra* note 130, at 276–77; *see also* Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 871 (1990).

190. Actually, it is the expectation of patent protection that provides the incentive. *See* Christopher A. Cotropia, *"After-Arising" Technologies and Tailoring Patent Scope*, 61 N.Y.U. ANN. SURV. AM. L. 151, 169–71 (2005).

191. *See* Gideon Parchomovsky & Peter Siegelman, *Towards an Integrated Theory of Intellectual Property*, 88 VA. L. REV. 1455, 1466–67 (2002). Parchomovsky and Siegelman state:

> [A]bsent legal protection, competitors would copy such works without incurring the initial costs of producing them . . . . [and, therefore,] [u]nauthorized reproduction would drive down the market price to the cost of copying, original authors and inventors would not be able to recover their expenditures on authorship and R&D, and, as a result, too few inventions and expressive works would be created."

Id. at 1467.

192. *See* Cargill, Inc. v. Canbra Foods, Ltd., 476 F.3d 1359, 1363 (Fed. Cir. 2007) (stating that "an affirmative misrepresentation of material fact" or "submitt[al of] false material information" can be inequitable conduct).

disclosure and misrepresentation.[193] As a result, patent attorneys are likely to exercise more care when drafting their communications with the USPTO. Patent attorneys ensure that everything discussed relevant to patentability is true and that nothing could be construed as a misrepresentation. This leads to a more accurate and readable public record, multiplying the benefits of a high quality patent application discussed above.

The current intent standard dampens the level of care required by the doctrine, however. The patent attorney must specifically intend to make a false statement or mislead the patent examiner; gross negligence is not enough.[194] However, misrepresentations are still considered material under inequitable conduct. Their appearance in an application or response to an office action will likely prompt at least an allegation of inequitable conduct. This potential exposure at least affects the behavior of some patent attorneys, making them more careful in what they write.[195]

IV. CURRENT INEQUITABLE CONDUCT DOCTRINE RESULTS IN OVERCOMPLIANCE

The inequitable conduct doctrine can only increase patent quality through compliance by patent applicants. The doctrine does this, like most other legal doctrines, by imposing certain legal and extra-legal costs on those who do not comply. In their current form, however, these costs are extremely high and instead of causing compliance, they prompt overcompliance.

---

193. *See* Purdue Pharma L.P. v. Endo Pharm. Inc., 438 F.3d 1123, 1132 (Fed. Cir. 2006) (finding material the use of the word "discovery" to characterize the invention while it was merely an "insight"—no test had actually been performed).

194. *See* Kingsdown Med. Consultants, Ltd. v. Hollister Inc., 863 F.2d 867, 872–73 (Fed. Cir. 1988) (en banc).

195. *See, e.g.*, Stephen K. Sullivan, *Drafting a Biotechnology Patent Specification*, *in* 16TH ANNUAL ADVANCED PATENT PROSECUTION WORKSHOP: CLAIM DRAFTING & AMENDMENT WRITING, at 135, 145 (PLI Patents, Copyrights, Trademarks, & Literary Prop., Course Handbook Series No. 9100, 2006) (instructing that, in light of *Hoffman-La Roche, Inc. v. Promega Corp.*, 323 F.3d 1354, 1363–64 (Fed. Cir. 2003), which found inequitable conduct "where prophetic examples were presented in the past tense, as if they had actually been performed," an attorney should "[b]e careful with word tense"). While it is true that inequitable conduct is likely to be alleged regardless of the patent attorney's actions, being more careful will have an impact on the number of false positives—improper findings of inequitable conduct.

Failure to comply with the doctrine renders the whole patent and, potentially, related patents, unenforceable.[196] A finding of inequitable conduct also exposes the patentholder to antitrust liability and liability for attorney fees.[197] The doctrine imposes specific costs on the patent attorney too, ranging from disciplinary action from the USPTO and the applicable state bar to malpractice liability to irreparable damage to the attorney's reputation.[198] All these costs are high in absolute terms and become even greater when compared to the low costs of overcompliance—simply submitting all information in one's possession to the USPTO, regardless of its materiality. This dramatic cost differential combined with uncertainty inherent in the inequitable conduct doctrine leads to overcompliance. This line of analysis is explored in detail below.

A. Breadth of Remedies Makes Non-Compliance Extremely Costly

Initiating a patent lawsuit exposes the patent holder to a range of liabilities. The patentee may, if unsuccessful, be saddled with the other side's attorney fees under the fee-shifting statute.[199] More significantly, the patentee may lose any of the asserted patent claims if they are found to be invalid. A final judgment of invalidity prevents the patentee from successfully asserting the now invalid claim against other infringers.[200] Thus, by asserting particular patent claims in a given lawsuit, the patentee is putting those claims at risk.[201] Patentees have to weigh the potential benefits of

---

196. *See* Star Scientific, Inc. v. R.J. Reynolds Tobacco Co., 537 F.3d 1357, 1365 (Fed. Cir. 2008) ("[T]he penalty for inequitable conduct is so severe, the loss of the entire patent even where every claim clearly meets every requirement of patentability.").

197. *See* Walker Process Equip., Inc. v. Food Mach. & Chem. Corp., 382 U.S. 172, 177 (1965) (discussing antitrust liability for asserting a patent procured through inequitable conduct); Brasseler, U.S.A. I, L.P. v. Stryker Sales Corp., 267 F.3d 1370, 1380 (Fed. Cir. 2001) (concluding that a finding of inequitable conduct results in an "exceptional" case, resulting in an award of attorney fees).

198. *See* Kelly Merkel, *How to Stump a Corporate Lawyer: Means of Effective Legal Risk Management for IP Counsel*, 1 J. LEGAL TECH. RISK MGMT. 1, 3 (2006) ("A finding of inequitable conduct can therefore open the door for suits by the client against the practitioner for breach of fiduciary duty and malpractice, even years after prosecution has ended.").

199. *See* 35 U.S.C. § 285 (2006) (awarding the "prevailing party" attorney fees in "exceptional cases").

200. *See* Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313, 330–31 (1971) (holding that once the claims of a patent are held invalid, the patent holder is collaterally estopped from enforcing the claim against another party).

201. *See* Cardinal Chem. Co. v. Morton Int'l, Inc., 508 U.S. 83, 102–03 (1993) (instructing courts to rule on invalidity regardless of the outcome on infringement).

enforcing their patent—monetary damages and an injunction[202]—with the risk of losing patent claims and paying attorney fees. A patentee may pay for high cost litigation with nothing to show for it—no finding of infringement or remedies because the patent is adjudged unenforceable.[203]

As compared to invalidity, the inequitable conduct doctrine places the patent holder in far less control over the downside of enforcing a patent claim. Invalidity affects only those asserted patent claims.[204] If a patentee does not want to risk the value in a particular claim, she simply does not assert it. Inequitable conduct, in contrast, causes the assertion of a single patent claim to expose the whole patent, and potentially all related patents, to a finding of unenforceability.[205] Sure, she still has control on a broader level as to what patent families are exposed. But her control is not nearly as fine as compared to her ability to cabin the impact of invalidity.

The costs resulting from a finding of inequitable conduct also include liability for attorney fees and exposure to antitrust liability. The patent statutes give courts the ability to award a successful party its attorney fees if the case is "exceptional."[206] A finding of inequitable conduct typically makes the case exceptional[207] and results in a fee award, which can reach

---

202. *See* 35 U.S.C. §§ 283, 284 (2006).

203. *See* AM. INTELLECTUAL PROP. LAW ASS'N, AIPLA REPORT OF THE ECONOMIC SURVEY I-90 (2007) [hereinafter AIPLA REPORT] (reporting that average cost taking a patent infringement case with less than one million dollars at risk through discovery being $461,000).

204. *See* DyStar Textilfarben GmbH & Co. Deutschland KG v. C.H. Patrick Co., 464 F.3d 1356, 1372 (Fed. Cir. 2006) (noting that invalidity is determined on a "claim-by-claim basis").

205. Fox Indus., Inc. v. Structural Pres. Sys., Inc., 922 F.2d 801, 803–804 (Fed. Cir. 1990) (stating that inequitable conduct "may render unenforceable all claims which eventually issue from the same or a related application"); *see also* Consol. Aluminum Corp. v. Foseco Int'l Ltd., 910 F.2d 804, 812 (Fed. Cir. 1990) (finding that the inequitable conduct during prosecution of one patent "permeated the prosecution of the other" patents-in-suit). Under certain circumstances, inequitable conduct will not spread. *See, e.g.*, Baxter Int'l, Inc. v. McGaw, Inc., 149 F.3d 1321, 1331–32 (Fed. Cir. 1998). The court in Baxter stated:

> [W]here the claims are subsequently separated from those tainted by inequitable conduct through a divisional application, and where the issued claims have no relation to the omitted prior art, the patent issued from the divisional application will not also be unenforceable due to inequitable conduct committed in the parent application.

Id. at 1332.

206. 35 U.S.C. § 285 (2006).

207. *See* Brasseler, U.S.A. I, L.P. v. Stryker Sales Corp., 267 F.3d 1370, 1380 (Fed. Cir. 2001) ("The prevailing party may prove the existence of an exceptional case by showing: inequitable conduct before the PTO . . . .").

well into the seven figure range.[208] The assertion of a patent obtained by inequitable conduct may also be subject to antitrust liability.[209] "If a patentee asserts a patent claim and the defendant can demonstrate the required fraud on the PTO, as well as show that 'the other elements necessary to a [Sherman Act] § 2 case are present,' the defendant-counterclaimant is entitled to treble damages under the antitrust laws."[210]

B. Doctrine's Specific Impact on Patent Attorneys Makes Non-Compliance Even More Costly

A finding of inequitable conduct does not directly result in personal liability for the patent attorney.[211] Nor does such a finding necessarily include a factual holding that the patent attorney was at fault.[212] Inequitable conduct can occur where the attorney disclosed all she knew, but one of the other parties under the duty intentionally failed to come forward with material information.[213]

However, the patent attorney is invariably at the center of any inequitable conduct inquiry. The patent attorney acts as the hub for the information flow from inventor and related parties to the USPTO. She assists in drafting the patent application and correspondence with the examiner. She typically signs all correspondence with the USPTO.[214] Her name also appears on the front of the issued patent.[215] Finally, the patent attorney is the one who best understands the legal obligations set forth by the inequitable conduct doctrine and usually communicates these obligations to the other relevant parties. So any non-disclosure, even if not her fault, has the patent attorney's fingerprints on it. As a result, the patent attorney is usually the first person noticed for deposition on the inequitable conduct issue and almost always mentioned by name in any inequitable conduct decision.[216]

---

208. *See* AIPLA REPORT, *supra* note 203, at 1-93.

209. *See* Dippin' Dots, Inc. v. Mosey, 476 F.3d 1337, 1346–47 (Fed. Cir. 2007).

210. Walker Process Equip., Inc. v. Food Mach. & Chem. Corp., 382 U.S. 172, 176–77 (1965).

211. *See, e.g.*, Star Scientific, Inc. v. R.J. Reynolds Tobacco Co., 537 F.3d 1357, 1365 (Fed. Cir. 2008) (explaining that the focus of the doctrine is rendering the patent at issue unenforceable).

212. *See* 37 C.F.R. § 1.56 (a), (c) (2008) (detailing the various individuals beyond the attorney who are under a duty to disclose).

213. *See id.*

214. *See* 37 C.F.R. § 1.34 (2008).

215. *See* 37 C.F.R. § 1.51(c)(v) (2008).

216. *See, e.g.*, McKesson Info. Solutions, Inc. v. Bridge Med., Inc., 487 F.3d 897, 903 (Fed. Cir. 2007) (referring to the prosecuting attorney by first and last name); Golden Hour Data Sys., Inc. v. emsCharts, Inc., No. 2:06 CV 381, 2009 WL 781334, *2 (E.D.

The doctrine truly puts "the person on trial, not the patent," and that person is the patent attorney.[217]

The doctrine can result in personal legal costs for the patent attorney involved. A failure to comply with doctrine can form the basis for a disciplinary action before the USPTO.[218] The patent attorney can also lose her license to practice before the USPTO.[219] The matter may be referred to her state bar, where the patent attorney may be disciplined or even lose her general license to practice law.[220] A judgment of inequitable conduct can also form the basis of a malpractice claim.[221]

The personal costs can also be extra-legal. Allegations of inequitable conduct implicate a patent attorney's professionalism and reputation in the legal community at large and before the USPTO, where she is a repeat player.[222] This reputation and personal liability exposure are so important to patent attorneys that some have even moved to personally intervene in inequitable conduct cases. Recently, in *Nisus Corp. v. Perma-Chink Systems, Inc.*, a patent attorney who prosecuted the patent at issue filed a motion to intervene in the patent infringement litigation and asked the district court to reconsider the conclusion of inequitable conduct.[223] He specifically challenged the court's "characteriz[ation] of his behavior in the court of the prosecution as constituting inequitable conduct."[224]

Tex. Mar. 23, 2009) (identifying both the prosecuting patent agent and supervising attorney by name).

217. Chisum, *supra* note 9, at 279; *see also* Aventis Pharma S.A. v. Amphastar Pharm., Inc., 525 F.3d 1334, 1349–50 (Fed. Cir. 2008) (Rader, J., dissenting) (noting that the "allegation of inequitable conduct opens new avenues of discovery" targeted at the patent attorney, not the patent).

218. See Edwin S. Flores & Sanford E. Warren, Jr., Inequitable Conduct, Fraud, and Your License to Practice Before the United States Patent and Trademark Office, 8 TEX. INTELL. PROP. L.J. 299, 314–15 (2000).

219. *Id.*

220. *Cf. id.*

221. *See* David Hricik, How Things Snowball: The Ethical Responsibilities and Liability Risks Arising from Representing a Single Client in Multiple Patent-Related Representations, 18 GEO. J. LEGAL ETHICS 421, 459 (2005).

222. *See* Migliorini, *supra* note 120, at 260 ("No client is worth the risk to one's personal integrity, reputation, and license to practice before the Bar" by committing inequitable conduct); *see also Aventis*, 525 F.3d at 1349–50 (Rader, J., dissenting) (indicating how allegations of inequitable conduct "impugns the integrity of the patentee, its counsel, and the patent itself").

223. 497 F.3d 1316 (Fed. Cir. 2007).

224. *Id.* at 1318.

C. High Cost of Non-Compliance, Uncertainty in Court Decisions, and Low Cost of Overcompliance Results in Overcompliance

Typically, law shapes behavior by making the costs of non-compliance outweigh any benefits.[225] Law relies on both legal and extra-legal costs to incentivize individual compliance.[226] As a result, rational, risk-neutral individuals do exactly what the legal rule requires to avoid engaging in behavior that is, on balance, detrimental to that individual. Applying this to the inequitable conduct doctrine, the doctrine uses certain legal and extra-legal costs to prompt patent applicants to disclose material information to the USPTO. Rational, risk-neutral patent applicants therefore respond by disclosing material information.[227]

Legal rules almost always have some inherent ambiguity, either because the scope of the rules is uncertain, or because the likelihood of enforcement is not absolute. Individuals may not know ex ante exactly what they must do to comply with a given doctrine. Inequitable conduct is no different. It, like most patent doctrines, has some ambiguities.[228] Determinations of whether a piece of information is material are difficult.[229] Materiality is a multi-step inquiry, involving the determination of each patent claim's meaning, analysis of the content of the information in question, and a judgment as to whether the information is relevant to issues of novelty, nonobviousness, or the disclosure requirements.[230] Whether the nec-

---

225. *See, e.g.*, Roger D. Blair & Thomas F. Cotter, *An Economic Analysis of Damages Rules in Intellectual Property Law*, 39 WM. & MARY L. REV. 1585, 1617–19 (1998) ("In order to deter infringement, we must have a set of rules that renders an infringement unprofitable.").

226. Legal costs are those costs, such as damage awards or injunctions that are imposed directly by the law—the remedial regime. Extra-legal costs are those costs, such as reputation or guilt, that do not derive directly from a statute or legal rule.

227. If an applicant is risk-adverse, they will overcomply even if what is required for exact compliance is clear. If an applicant is risk-seeking, the opposite is true—undercompliance. *See* James Gibson, *Doctrinal Feedback and (Un)reasonable Care*, 94 VA. L. REV. 1641, 1651 (2008) (noting that while speed limits provide bright-line liability rule, "[a] risk-averse driver will drive more slowly, and a risk-seeking driver will drive more quickly").

228. The focus is on the ambiguity of the rule, not its enforcement. The patent holder control its enforcement, opening herself up to such a defense when asserting her patent.

229. *See* Alpa Gandhi, *The Fate of the Rule 56 Materiality Standard in the Inequitable-Conduct Inquiry*, 33 AIPLA Q.J. 125, 127–28 (2005) (describing the uncertainty in the materiality standard).

230. Even the first step of this process has been empirically proven to be incredibly uncertain. Christian A. Chu, *Empirical Analysis of the Federal Circuit's Claim Construction Trends*, 16 BERKELEY TECH. L.J. 1075, 1097–99 (2001) (finding a 50% reversal rate for claim interpretations).

essary subjective intent is present is often ambiguous, particularly given that intent is proven in court through circumstantial evidence.[231] An applicant cannot predict with absolute certainty how a court will decide these requirements given their fact-dependency.

Patent attorneys, and others under the duty to disclose, are faced with an array of choices as to how to comply. For example, what type of information should they submit to the USPTO? And under what circumstances should they submit such information? They essentially must make a choice as to whether they err on the side of undercomplying or overcomplying—either of which, depending on the distribution of uncertainty around the legal rule, has a certain probability of avoiding liability.[232] Undercompliance is typically chosen where there are, on net, substantial benefits to undercompliance that outweigh the risk and impact of being found liable.[233] In contrast, an individual chooses to overcomply where the costs of overcompliance are small compared to the costs of being found liable.[234]

As previously discussed, the inequitable conduct doctrine makes a finding of non-compliance extremely costly.[235] The doctrine extracts both legal and extra-legal costs on both the patent holder and the patent attorney. In comparison, the costs associated with overcompliance are minimal. The most common method of overcomplying under the current legal regime is to submit everything of even remote relevance in one's possession to the USPTO.[236] Even if the information is not material to the claimed invention, disclosure absolves any potential violation of the doctrine.[237]

---

231. *See* D. Ward Hobson Jr., Reforming the Patent System: A Closer Look at Proposed Legislation, 3 OKLA. J. L. & TECH. 29 (2006).

232. John E. Calfee & Richard Craswell, *Some Effects of Uncertainty on Compliance with Legal Standards*, 70 VA. L. REV. 965, 971–74 (1984) (demonstrating the different distribution of uncertainty and describing the possible causes of over and undercompliance).

233. *Id.* at 981.

234. *Id.* at 981–82.

235. *See supra* Part IV.

236. *See* Thomas, *supra* note 17, at 315 ("Where the applicant is already well informed of the prior art, the specter of inequitable conduct too often causes applicants to submit virtually every reference of which they are aware.").

237. Moreover, the disclosure of non-material or cumulative references imposes no penalty on the disclosing party. *See* Molins PLC v. Textron, Inc., 48 F.3d 1172, 1184 (Fed. Cir. 1995) (burying a material reference in a voluminous submission of information is not actionable unless there is specific intent to hide the reference).

The common mantra is "when in doubt, disclose."[238] Even Federal Circuit case law, particularly recent case law, gives this advice.[239]

But every quanta of information submitted has its costs. The patent applicant must submit an IDS with the information, which must include a listing of each reference being submitted and identify for each reference, where relevant, the "publisher, author (if any), title, relevant pages of the publication, date, and place of publication."[240] The applicant is also required to submit a legible copy of any reference that is not another U.S. patent or published U.S. patent application.[241] If the applicant needs to disclose a piece of art that is in a foreign language, the applicant must also submit a "concise explanation of the relevance" of the foreign language reference.[242] Depending on the timing of the information's discovery, the applicant may have to pay the fees associated with continuing prosecution to allow the examiner to consider the newly submitted information.[243]

These costs, however, are minuscule compared to losing the enforceability of a valid patent, or possibly a whole family of valid patents. And when viewed from the patent attorney's perspective, overcompliance looks even more inviting. The patent attorney, while getting some satisfaction (and presumably repeat business) from successfully obtaining a patent, obtains no personal gain from the issuance of the patent. In contrast, she has significant concerns that getting caught up in an inequitable conduct claim will damage her livelihood. Added to this is the fact that overcompliance generates more legal fees. The attorney gets to charge her client for the time required to submit the additional information and continue prosecution if necessary. So, even if overcompliance becomes marginally

---

238. *See* Thomas C. Fiala & Jon E. Wright, *Preparing and Prosecuting a Patent That Holds Up in Litigation*, *in* PATENT LITIGATION 2006, at 515, 547 (PLI Patents, Copyrights, Trademarks, & Literary Prop., Course Handbook Series No. 9001, 2006) ("If it is unclear whether information is prior art, whether it is 'material', or whether it is cumulative to information already submitted to the USPTO, the information should be disclosed so that the examiner can make the determination.").

239. *See* Praxair, Inc. v. ATMI, Inc., 543 F.3d 1306, 1313–14 (Fed. Cir. 2008) (stating that an applicant must have a "credible good faith explanation for the withholding"); Critikon, Inc. v. Becton Dickinson Vascular Access, Inc., 120 F.3d 1253, 1257 (Fed. Cir. 1997) (noting that doubts concerning whether the information is material should be resolved in favor of disclosure).

240. 37 C.F.R. § 1.98 (2008).

241. § 1.98(2).

242. § 1.98(3)(i).

243. The costs for filing a continuation in these instances, called a Request for Continued Examination ("RCE"), includes a filing fee and accompanying attorney billable hours to put the filing together. *See* 35 U.S.C. § 132 (2006); 37 C.F.R. § 1.98; 37 C.F.R. § 1.114 (2008).

expensive, patent attorneys may still push clients to overcomply because of their self-interest in such a course of action.

V. OVERCOMPLIANCE CAUSED BY THE INEQUITABLE CONDUCT DOCTRINE REDUCES PATENT QUALITY

Overcompliance puts any improvement in patent quality created by the inequitable conduct doctrine at risk. It overloads the examiner with information that, in most instances, is immaterial. The examiner, with an extremely small amount of time to examine highly technical subject matter, does not process all of the submitted information or ignores it altogether, erasing any quality gains. In fact, the doctrine may end up doing more harm than good. Overcompliance can so stress the examiner as to impair her ability to make a sound decision based on the information she does process. Overcompliance further prices inventors out of the patent system, causing its own set of societal harms. These harms from overcompliance are discussed below.

A. Overcompliance Causes Detrimental Information Overload

The inequitable conduct doctrine is focused on getting quality information before the USPTO. The doctrine requires only material information be submitted. The patent applicant is not required to submit "information which is not material to the patentability of any existing claim."[244] The doctrine places no weight on the quantity of information placed before the patent examiner. In fact, the doctrine invites the applicant to thin her submissions by not requiring the submittal of cumulative information.[245]

However, as established above, the doctrine incentivizes the patent applicant to err on the side of quantity.[246] Applicants make the safe play and overcomply. They disclose all of the information within their possession that is remotely relevant to the claimed subject matter. The doctrine, therefore, causes examiners to receive additional quantities of information that are increasingly immaterial to the task at hand—determining patentability.

The patent examiner, with this additional information generated by overcompliance, can experience information overload. Information overload occurs when a decision-maker cannot naturally process the information in their possession in an allotted time without a high likelihood of

244. 37 C.F.R. § 1.56(a) (2008).
245. *See supra* Section II.B.
246. *See supra* Part IV.

making mistakes.[247] The chances of overload are particularly high when the information is highly technical or complex.[248] Examiners are overworked, with an increasing number of patent applications to examine in an ever-decreasing amount of time.[249] And the information they must process—the application and prior art—is technical by definition. Adding information submitted by the applicant to the mix, particularly large amounts of information from those who overcomply, plays a significant part in overloading the examiner.[250]

Information overload can negate any benefit in patent quality gained by the inequitable conduct doctrine. When overloaded, an individual has difficulty identifying information relevant to the decision-making task at hand.[251] An individual may overlook the most critical information.[252] The overloaded examiner must choose where to allocate her finite examining time. She may have to choose which of the submitted references she will read.[253] In the overload situation, the submitted information becomes increasingly immaterial, meaning the examiner will waste at least some of her time reading non-material information.[254] The bigger the haystack, the more lost a needle becomes.

Information overload can even cause the examiner to become so overwhelmed that she does not even attempt to sift through the applicant's submissions.[255] She ignores them completely. The benefits of the additional, relevant information the doctrine generates are lost in the sea of information.

---

247. Naresh K. Malhotra, Reflections on the Information Overload Paradigm in Consumer Decision Making, 10 J. CONSUMER RES. 436, 437 (1984).

248. *See* S.C. Schneider, *Information Overload: Causes and Consequences*, 7 HUMAN SYS. MGMT. 143, 144 (1987).

249. *See supra* Section III.A.1.

250. *See, e.g.*, Noveck, *supra* note 140, at 148-49 (noting that successful examination reforms need to include mechanism to prevent "overwhelming the patent examiner with information").

251. *See* Paul A. Herbig & Hugh Kramer, *The Effect of Information Overload on the Innovation Choice Process*, 11 J. OF CONSUMER MARKETING 45, 45 (1994).

252. *Id.*

253. *See* Thomas, *supra* note 17, at 315 ("Coupled with the severe time constraints facing the examining corps, this overload of information often allows no more than a cursory review of all but a few references that initially appear the most promising.").

254. The concept that examiners have a definite saturation point is further supported by recent empirical research that found the likelihood of receiving a rejection plateaus at twenty references. *See* Crouch, *supra* note 164 (finding the percentage likelihood hovering around 40% once twenty references is reached).

255. *Id.*

The harm from information overload can go beyond wiping out the doctrine's quality gains. In an attempt to process all of the information, the examiner simplifies her processing strategy.[256] This results in poorer decisions because fidelity is lost across the board—none of the information is properly processed.[257] She loses her ability to identify the relationship between details and her overall perspective on the decision at hand.[258] She becomes stressed, confused, and generally cognitively strained, impairing her ability to think analytically.[259] More becomes less.

Empirical studies indicate that this analysis holds true even if the additional information is as material as that already submitted.[260] That is, the increase in quantity still overwhelms the decision-maker even if the additional information is of high quality. The materiality of the additional information is irrelevant to the information overload scenario. The overwhelming volume of the information degrades the examiner's ability to reach a proper decision.

B. Results in Socially-Wasteful Costs

As previously stated, the costs of overcompliance are small when compared to the costs of being found non-compliant. To the patentee and the attorney, spending a few thousand dollars, even tens of thousands of dollars, and delaying the issuance of the patent is miniscule compared to losing a full family of patents, paying attorney fees, facing exposure to antitrust liability, risking possible bar discipline, and so on. However, this is the private, internal cost-benefit analysis.

The answer to the question of costs is very different when looked at in terms of the big picture. That is, is this spending beneficial to the public? The answer is no, given that overcompliance artificially increases the price of patent procurement, and the extra dollars spent going beyond what is required under the inequitable conduct doctrine add nothing to the quality of the patent examination. In fact, as discussed above, the submission of additional information, particularly immaterial information, actually ham-

---

256. *See* Naresh K. Malhotra, *Information Load and Consumer Decision Making*, 8 J. CONSUMER RES. 419, 427 (1982) (noting that information overload causes individuals to "adopt simplifying information-processing strategies").

257. *See* Kevin Lane Keller & Richard Staelin, *Effects of Quality and Quantity of Information on Decision Effectiveness*, 14 J. CONSUMER RES. 200, 212 (1987) (concluding that overload "degrade[s] choice accuracy").

258. *See* Schneider, *supra* note 248, at 145.

259. *See* Malhotra, *supra* note 256, at 427.

260. *See* Keller & Staelin, *supra* note 257, at 212.

pers the examination process by creating information overload.[261] Over-compliance is socially wasteful spending.

These additional costs can also create another harm—pricing out potential patentees. The costs of compliance are not significant. In most cases, they may constitute, at most, one percent of the total cost of obtaining a patent.[262] However, price tolerances for patenting can be extremely sensitive, especially for individual inventors or small companies.[263] Each increase in the cost of patenting can deter these would-be inventors from inventing altogether, or, prompt them to choose trade secret protection in lieu of patenting.[264] Both of these options are disadvantageous, possibly robbing society of the next great invention or hiding the details of that invention from the general public.

One option for these individuals, for which compliance is too costly, is to undercomply.[265] But given that the patent attorney has significant individual interests at stake, the attorney is unlikely to play along. In fact, the professional advice to patent attorneys is to "avoid being pressured by clients to compromise [their] ethical duties."[266] Furthermore, the cost of compliance with the doctrine is likely not transparent to the cost-sensitive applicant. The cost is simply included in initial quote for the cost of getting a patent. The small inventor has no practical choice to forgo compliance.

VI. USING THIS FRAMEWORK TO REFORM THE INEQUITABLE CONDUCT DOCTRINE

The benefits of constructing a modern framework for the inequitable conduct doctrine are two-fold. First, the framework identifies ways in which the doctrine can be beneficial and detrimental. From there, reforms can be targeted, attempting to maximize the positive aspects of the doctrine and minimize the negative ones. Second, the framework provides a workable model upon which current concerns and suggested reforms can be vetted. And given that the framework established is utilitarian based, it

---

261. *See supra* Section V.A.

262. *See* AIPLA REPORT, *supra* note 203, at 1-78 (finding the average cost of obtaining a patent between $6,600-$15,000).

263. Patent law has recognized this fact, establishing a separate fee schedule for "small entit[ies]." *See, e.g.*, 37 C.F.R. § 1.16(a) (2008) (setting forth lesser filing fees for small entities).

264. *See, e.g.*, Wagner, *supra* note 143, at 236–37 (recognizing that increasing the costs of prosecution "decreases the incentives produced by the patent system").

265. *See* Calfee & Craswell, *supra* note 232, at 981–82.

266. Migliorini, *supra* note 120, at 260.

is much easier to address recent criticisms, which all have a utilitarian bent.

This Part uses the analysis previously performed to do both of these things. Reforms are initially suggested to remedy the overcompliance currently provoked by the doctrine. The extent of the remedies for violating the doctrine needs to be reduced, a specific intent standard separate from materiality must be maintained, and the submission of immaterial and cumulative information discouraged. The positive aspects of the doctrine should, however, not be lost, and so, the materiality standard must remain broad.

Next, the Article's framework is applied to the two most common criticisms regarding the doctrine. The duties under the doctrine should not be expanded to include a duty to search or provide relevancy statements. Such duties are likely to overload the examiner, price inventors out of the patent system, shift the burdens of examination away from a low cost provider, and destroy the benefits of independent review. Finally, if the reforms proposed in this Part are adopted, a reduction in litigation costs follows. Thus, there is no need for any specific reforms to address this perceived problem.

A. Reducing the Likelihood of Overcompliance

The cost of non-compliance can be reduced, the ambiguity surrounding the doctrine can be minimized, or the costs of overcompliance can be increased. All three options are suggested below.

1. Minimize the Remedy

The current remedy is incredibly severe—truly an "atomic bomb."[267] One way of reducing the amount of over compensation is to minimize the costs associated with non-compliance. It is the high cost of not complying that, in part, drives applicants to overcomply.[268] The question is to what extent should the costs of non-compliance be reduced. How far should the available legal remedies and associated legal and extra-legal costs be reigned in?

A good place to start is to tie the legal remedy with the harm non-disclosure does to patent quality. Failure to disclose material information hampers the examination of those patent claims to which the material is relevant.[269] The remedy should be adjusted accordingly. No longer should

---

267. Aventis Pharma S.A. v. Amphastar Pharm., Inc., 525 F.3d 1334, 1349 (Fed. Cir. 2008) (Rader, J., dissenting).

268. *See supra* Part IV.

269. *See supra* Part III.

a finding of inequitable conduct render the whole patent, and possibly related patents, unenforceable. The remedy should mimic a finding of invalidity—only those claims to which the undisclosed information is material should be rendered unenforceable. The patent holder can then control her exposure by subjecting only asserted claims to a finding of unenforceability.

The legal remedies could be reduced further, taking the form of a monetary remedy, for example. A finding of inequitable conduct could result in a fine or the damage award for infringement could be reduced. The problem with swinging this far in the other direction is that, by taking the patent out of harms way, the applicant may undercomply. The fine or potential reduction in damages could simply be folded in with the cost of enforcing the patent. And to make monetary damages effective, the amount would have to be adjusted in light of the potential upside—monetary and injunctive remedies—to the patent holder benefiting from successful enforcement. This introduces uncertainty into the remedy regime, which makes it even more difficult to predict the extent to which applicants will comply.[270] Adjusting the monetary remedy is an imperfect way of getting at the patent's value to the patent holder. If that is the goal, the simple solution is to find the asserted claims unenforceable.

A final possible remedy is to give district courts the discretion in determining the remedy.[271] The court can vary the equitable relief accordingly, from finding the whole family of patents unenforceable to denying injunctive relief. Discretion allows the district court to fine-tune the remedy to the facts of a particular case. But *ex ante*, when the applicant is trying to determine how to comply, the applicant has no idea what remedy a judge will choose. This introduces uncertainties of its own, which increases the variation in the levels of compliance. Furthermore, if the extreme remedy is still in play—unenforceability of a group of patents—rational applicants are still likely to overcomply.

2. *Maintain a Specific, Independent Standard for Intent*

The high costs unique to the patent attorney—potential disbarment, malpractice liability, and damage to reputation—may cause overcompliance to continue even if the remedies are reduced. The best way to reduce these attorney-specific costs is not through just changing the costs of noncompliance, but reducing the uncertainty in the doctrine, particularly as it pertains to the attorney. As the target—how to comply with the doctrine—

---

270. *Cf.* Calfee & Craswell, *supra* note 232, at 971–72.

271. This change was proposed in the 2007 Patent Reform Act. *See* Patent Reform Act of 2007, H.R. 1908, 110th Cong. § 12(c)(4) (2007).

becomes clearer, it becomes easier for patent applicants, and their attorneys, to tune their response to actual compliance.[272]

The intent requirement in the inequitable conduct doctrine is the best candidate to reduce ambiguity. The materiality requirement is too complex for any reforms to provide much certainty because the standard for materiality *and* the underlying concept of patentability would need to be cleared up.[273] Intent, in contrast, is a familiar doctrine, common across multiple legal fields. The clearer the standard of intent under the inequitable conduct doctrine *ex ante,* the easier it will be for patent attorneys to feel confident that they will be found in compliance.

The "bleed through" from the materiality finding is responsible for most of the intent doctrine's current ambiguity. The recent conflict at the Federal Circuit over the intent requirement is evidence of this ambiguity.[274] A finding of high materiality—the information was very material to patentability—is used as conclusive circumstantial evidence that the applicant intended to deceive the patent examiner.[275] This type of analysis reduces the inequitable conduct inquiry into nothing more than a determination of materiality. It also makes it difficult for applicants, and their attorneys, to be confident that they are free from liability if they actually believe something is not material or if they unintentionally overlook information in their possession. Even if these facts are true, and they did not intend to deceive the USPTO, inequitable conduct will likely be found because of bleed through.

---

272. *See* Calfee & Craswell, *supra* note 232, at 971 (showing how the distribution of possible responses grows narrower as certainty reduces, minimizes the magnitude of possible overcompliance).

273. For example, the standard for nonobviousness is very unclear after the Supreme Court's recent decision in *KSR*. *See* Tun-Jen Chiang, *A Cost-Benefit Approach to Patent Obviousness*, 82 ST. JOHN'S L. REV. 39, 53–54 (2008).

274. *Compare* Star Scientific, Inc. v. R.J. Reynolds Tobacco Co., 537 F.3d 1357, 1366–67 (Fed. Cir. 2008) (stating that inferences of intent should go in the favor of the applicant and no finding of inequitable conduct), *with* Praxair, Inc. v. ATMI, Inc., 543 F.3d 1306, 1313–14 (Fed. Cir. 2008) (explaining that it's the applicant's burden to rebut an inference of intent with affirmative, credible evidence of good faith).

Judge Linn, in a recent concurrence, called for an en banc review of the intent standard in order to clarify the standard and identify acceptable evidence to prove intent. *See Larson Mfg.*, 2009 WL 691322, at *20.

275. Purdue Pharma L.P. v. Endo Pharm. Inc., 438 F.3d 1123, 1133–35 (Fed. Cir. 2006) ("[A] patentee facing a high level of materiality and clear proof that it knew or should have known of that materiality, can expect to find it difficult to establish 'subjective good faith' sufficient to prevent the drawing of an inference of intent to mislead.") (quoting Critikon, Inc. v. Becton Dickinson Vascular Access, Inc., 120 F.3d 1253, 1257 (Fed. Cir. 1997)).

This ambiguity leads to overcompliance. Applicants second-guess themselves, submitting information even if they personally believe the information is not material.

To resolve this ambiguity and make the lack of intent a true safe haven, a specific intent standard that is distinct from materiality needs to be adopted. The 2007 Patent Reform Act contained such a provision.[276] With such a requirement, patent applicants, and particularly attorneys, can better tailor their conduct to meet the doctrine's requirements. If they do not intend to deceive the USPTO—they truly believe that undisclosed information is not material, for example—they can be pretty sure they will not be found liable.

This solution does not create absolute certainty. Circumstantial evidence is still available to establish intent, and it should be. If the doctrine required direct evidence, it would be near impossible to establish inequitable conduct, resulting in massive undercompliance.[277] However, removing the ability of accused infringers to piggyback an intent finding based on materiality still reduces a significant amount of the ambiguity in the doctrine.[278] The more certain a path towards compliance, the less one overcomplies. Establishing a truly independent intent standard goes a long way in providing that certainty.

3. Prohibit the Submission of Cumulative and Non-Material Art

Even with a reduction in costs and a clearer legal standard, the incentive to overcomply likely remains. This incentive does not come solely from fear of the costs of non-compliance. Instead, it is also driven by the fact that overcompliance—submitting all information in the applicant's possession without reviewing for materiality—can still be the cheapest way to comply. Reading and evaluating a reference is the most attorney intensive, and thus expensive, part of submitting information to the USPTO. Thus, even if costs of non-compliance are reduced, applicants may still overcomply because it is the least costly type of compliance. Information overload, therefore, continues.

---

276. *See* H.R. 1908 § 5(c)(3) (2007) ("[S]pecific facts beyond materiality of the information misrepresented or not disclosed must be proven that establish the intent of the person to mislead or deceive the examiner by the actions of that person.").

277. *See* Hoffman-La Roche Inc. v. Lemmon Co., 906 F.2d 684, 688 (Fed. Cir. 1990) ("[I]ntent usually can only be found as a matter of inference from circumstantial evidence.").

278. Adoption of a presumption of no intent and the resolution of multiple inferences in favor of a finding of no intent is a move in the right direction. The Federal Circuit's recent articulation of the intent standard in Star Scientific follows such an approach. *See Star Scientific*, 537 F.3d at 1366–67.

Currently, an applicant who overloads the USPTO with immaterial or cumulative information does not commit inequitable conduct.[279] Even if the large volume of submissions effectively "buries" a particularly material reference, the applicant does not per se commit inequitable conduct.[280] Some district courts have used such activity as circumstantial evidence of intent.[281] But none, yet, have gone any further.

The potential solution is two fold. First, the intentional submission of immaterial or cumulative information should be actionable under the doctrine. This addresses the most egregious overloaders—those who are truly trying to bury the examiner with information they know is irrelevant. This does not, however, address those who are simply grossly negligent in their submission methodology by not reading references, or barely reading them, and then submitting them.

The second part of the solution is to actively enforce existing USPTO disciplinary rules that require applicants to read information before they submit it to the office. Rule 10.18(b)(2) requires patent attorneys to make an "inquiry reasonable under the circumstances" as to why a paper is submitted.[282] The rule requires the patent attorney to represent that "[t]he paper is not being presented for any improper purpose, such as to harass someone or to cause unnecessary delay or needless increase in the cost of prosecution before the [USPTO]."[283] Failure to comply with this rule risks the validity of the patent and sanctions against the attorney.[284] If this rule were actively enforced, it would incentivize applicants to read information before submitting it, with an eye toward not overloading the USPTO. Enforcement of this rule would temper the amount of low-cost overcompliance.[285]

---

279. *See supra* notes 226-227.

280. *See* Molins PLC v. Textron, Inc., 48 F.3d 1172, 1184 (Fed. Cir. 1995) (noting that the court must assume the examiner considered the submitted information).

281. *See, e.g.*, Penn Yan Boats, Inc. v. Sea Lark Boats, Inc., 359 F. Supp. 948, 964–65 (S.D. Fla. 1972).

282. 37 C.F.R. § 10.18(b)(2) (2008).

283. § 10.18(b)(2)(i).

284. 37 C.F.R. § 10.18(c) (2008).

285. The pending IDS rules, which require the filing of an Examination Support Doctrine ("EDS") if more then a particular number of pieces of prior art is filed is another possible avenue of reducing overcompliance. *See* Changes To Information Disclosure Statement Requirements and Other Related Matters, 71 Fed. Reg. 38,808 (July 10, 2006) (to be codified at 37 C.F.R. pt. 1). The ESD requires the applicant to identify the relevance of the submitted art to the pending claims, and the cost of such a doctrine arguably deters the oversubmission of art. However, this limitation is, in a sense, random, tied to a particular number. The "proper" number of references to submit varies with the technological area and the type of art in the applicant's possession in relationship to the claims.

B. Maintaining an Independent and Broad Materiality Standard

The inequitable conduct doctrine needs to maintain its positive effects on patent quality as well. To do this, the standard for materiality must stay independent of the standard for invalidity. That is, "materiality" need not require that the withheld information actually render a claim invalid.[286]

Requiring the submission of information that establishes only a prima facie case of invalidity, but does not make a conclusive case, broadens the patent quality gains under the doctrine. This standard includes information that is extremely relevant to the patent examination, but still broad enough to create spillover benefits. The more information required to be submitted, the larger the knowledge base of the patent attorney and greater the benefits from this increased knowledge.[287] This broad information base also increases the second-order information production benefits to examination. The examiner gets more information that may lead to additional relevant information or thinking that, in turn, produces a better examination.[288] However, this information base would not be so large and its relevancy not so tenuous as to significantly increase the likelihood of information overload.

In addition, if materiality were limited to claim-invalidating information, the inequitable conduct doctrine would become redundant during litigation. A finding of inequitable conduct may have secondary, legal cost effects on the patent holder. But once the claim is found invalid, the damage to that claim is done—the claim is no longer valid and thus cannot be enforced.[289] The value added by the doctrine is minimized greatly by equating the materiality standard with validity.

C. Avoiding Expansion of the Duties Governed by the Doctrine

Some critics question whether the duties governed by the doctrine should be expanded to include a duty to search and provide relevancy statements about submitted information. The 2007 version of the patent reform legislation contained a provision that gives the USPTO authority to require an applicant to do a search and inform the USPTO as to how the

---

There is no magical number of references where the cost benefit of providing them to the USPTO becomes negative.

286. *See, e.g.*, Patent Reform Act of 2005, H.R. 2795, 109th Cong. § 136(c)(2) (2005).

287. *See supra* Section III.C.1.

288. *See supra* Section III.B.1.

289. 35 U.S.C. § 282 (2006).

application is patentable in light of the search results.[290] The pending IDS rules proposed by the USPTO require relevancy statements in certain circumstances—again asking the applicant to link the submitted information to the application's claims.[291]

Expanding the applicant's duties to include either of these—searching or relevancy statements—would dramatically increase the cost of compliance. To perform a search in-house or request a search from an outside firm costs between $2,000 to $3,500.[292] Relevancy statements, which require the applicant to identify the relevant portions of the submitted information and distinguish the disclosed invention, are even more costly, ranging from $12,250 to $20,000.[293] These costs become more significant when compared to the typical cost of patent prosecution; they dwarf the cost of preparing and filing a typical patent application, which ranges from $6,600 to $15,000.[294] The ambiguous nature of these duties—how many databases to search, whether the search needs to be updated if new information is discovered during prosecution, whether new searches are required if the claims are amended—magnify these costs.[295] When the tendency to overcomply because of the legal and extra-legal costs exacted by the inequitable conduct doctrine is added, the costs of expanded duties become even higher.

These high costs could temper the current environment of overcompliance by making excessive disclosure cost-prohibitive.[296] This can reduce the problem of information overload. The additional duties can also magnify the patent quality effects of the inequitable conduct doctrine. The patent attorney would be exposed to more relevant information, making her

---

290. *See* Patent Reform Act of 2007, H.R. 1908, 110th Cong. § 11 (2007). The 2009 version of the legislation does not modify the inequitable conduct doctrine. *See* Patent Reform Act of 2009, S. 515, 111th Cong. (2009).

291. *See* Changes To Information Disclosure Statement Requirements and Other Related Matters, 71 Fed. Reg. at 38,808.

292. AIPLA REPORT, *supra* note 203, at 1-82.

293. *Id.* at 1-83 (describing the costs for a validity opinion). Relevancy statements, like those proposed in § 11 or 71 Fed. Reg. 38,808 are essentially requests for a validity opinion. That is, they ask the applicant to explain to the USPTO why the application patentable over the discovered information.

294. *Id.* at 1-78.

295. *See* Tafas v. Dudas, 511 F. Supp. 2d 652, 667–68 (E.D. Va. 2007) (noting the ambiguities inherent in such duties possibly rise to the level of being unconstitutionally vague). The Federal Circuit affirmed-in-part, vacated-in-part, and remanded the district court's decision in *Tafas*, but specifically did not address the potentially vagueness of the rules. Tafas v. Doll, 559 F.3d 1345, 1364-65 (Fed. Cir. 2009).

296. *See* Calfee & Craswell, *supra* note 232, at 981–82.

think more about the patentability of the application, and increasing the amount of quality information before the examiner.

These gains are unlikely to be realized, however. First, even normal compliance with these additional duties is likely to overload the examiner. While these duties do provide a plethora of new information on the examiner, there is no mechanism to give her additional time to process it. The new information does not only come in the form of new references, but analysis by the patent applicant if relevancy statements are also required.

The high costs of compliance will likely not cause the applicant to undercomply with the doctrine. The patent attorney has a personal interest in complying that goes beyond a single client's interest in getting a patent at a low cost. Again, patent attorneys are advised to not "be[] pressured by clients to comprise [their] ethical duties."[297] This means that applicants will be forced to pay the high fees to comply or forgo patenting altogether. By essentially doubling the cost of obtaining a patent, the expansion of duties may price potential inventors out of the system altogether, which either deters the formal invention process or pushes it outside the patent system and into the realm of trade secret.[298]

Finally, making the applicant essentially perform the examiner's job—searching and textually analyzing patentability based on the search—is socially inefficient. Having applicants evaluate and produce information already in their possession makes sense because they are the lowest cost provider even if they are not good at analyzing patentability.[299] But, having an applicant conduct searches, something in which she is not necessarily an expert, is a poor use of resources. Searching, and in particular searching the invention's field of technology, is the examiner's profession. If more searching is required, the examiner is the one to do it.

Maintaining the independence and second-review benefits of examination is also a goal.[300] An examiner is unlikely to do more than verify the applicant's patentability analysis under these new expanded duties. New examiners will not have the opportunity to gain search skills in a patent system where searching is done by the applicant, and independent review becomes even more of a fallacy.

---

297. *See supra* note 245.

298. *See supra* note 243.

299. *See supra* Section III.B.1.

300. *See, e.g.*, Craig Allen Nard & John F. Duffy, *Rethinking Patent Law's Uniformity Principle*, 101 NW. U. L. REV. 1619, 1627–37 (2007) (describing the benefits and disadvantages of decentralized decision making).

D. Reducing Litigation Costs as a Result of the Proposed Reforms

A final concern regarding the doctrine that needs to be addressed is litigation costs. Critics assert that the doctrine is alleged too often and too costly to litigate.[301] The reforms already mentioned, while not eliminating the litigation costs entirely, greatly reduce them. Thus, no additional reform is needed to address this problem.

Most of the reforms aimed at reducing overcompliance also make the inequitable conduct defense less attractive to assert. The breadth of remedies would be reduced, with only individual claims being exposed to unenforceability. Specific, independent evidence of intent would be required to prove inequitable conduct. Both of these changes weaken the doctrine, making it harder to prove and the rewards less tantalizing. This would decrease the rate of assertion.

Declining to expand the duties under the doctrine would also keep litigation costs down. The more theories of liability under the doctrine, the more opportunities for an alleged infringer to assert the defense of inequitable conduct. As the duties under the doctrine increase, the easier it becomes for a defendant to find at least one plausible theory of inequitable conduct. The defendant is also able to keep the claim alive longer during litigation. By keeping the duty focused on information within the applicant's possession, the doctrine does not open up new doors through which the defense can be alleged.

Some of the reforms could increase litigation costs. If materiality were tied to the question of validity, the doctrine would no longer be an independent litigation tool. It would become infrequently litigated because of the need to prove invalidity as a predicate. So maintaining the distinction between materiality and validity does forgo one option for saving litigation costs. The addition of a theory of inequitable conduct based on intentional submission of immaterial or cumulative information will likely spark some additional allegations of inequitable conduct. As the liability theories expand, so does the room for the doctrine's assertion.

On net, however, these reforms would constrict the scope of the doctrine, thus litigation costs overall will decrease. In addition, this concern may not sit within the broader utilitarian goal of patent law. Gains in patent quality caused by the suggested reforms would likely outweigh the negative impact from the already high cost of litigating the doctrine. The breadth of positive impact is so much greater—affecting every patent and

301. *See supra* Section II.C.2.

patent examination—than the small percentage of patents that are actually litigated.[302]

VII. CONCLUSION

One of the most pressing questions in the U.S. patent system is how responsibilities should be shared between the inventor and the USPTO when examining a patent application. The inventor wants to externalize costs by burdening the USPTO with the majority of the work and the USPTO, similarly, wants the inventor to internalize as much of the costs as possible. The inequitable conduct doctrine, which governs the inventor's duties during patent examination, sits at the center of this tension. The doctrine addresses when the applicant needs to assist in examination by providing information to the USPTO. The question has always been how much information needs to be provided.

This Article exposes the complexity of this problem by constructing a conceptual framework by which to measure the inequitable conduct doctrine's impact on the patent system. The doctrine can have a very positive impact on the system, by improving patent quality, but can also have tremendous negative impacts, by hindering examination and denying access to the incentive to invent. Knowing how the doctrine impacts the utilitarian goals of the patent system is crucial when determining how to tune the inequitable conduct doctrine. But understanding these dynamics tells even more about the patent application process and how examiners, applicants, and potential inventors are impacted by shifts in the cost-sharing of patent examination. This fundamental understanding, and the specific reforms it suggests, can help not only improve the inequitable conduct doctrine, but also the entire patent system.

---

302. *See* Mark A. Lemley & Carl Shapiro, *Probabilistic Patents*, 19 J. ECON. PERSP. 75, 75 (2005) (noting that only 1.5 % of issued patents are ever asserted and only 0.1% go to trial).

PRIVILEGE-WISE AND PATENT (AND TRADE SECRET) FOOLISH? HOW THE COURTS' MISAPPLICATION OF THE MILITARY AND STATE SECRETS PRIVILEGE VIOLATES THE CONSTITUTION AND ENDANGERS NATIONAL SECURITY

By Davida H. Isaacs[†] *and Robert M. Farley*[‡]

TABLE OF CONTENTS

[†] Associate Professor of Law, Salmon P. Chase College of Law, Northern Kentucky University.

[‡] Assistant Professor of National Security, Patterson School of Diplomacy and International Commerce, University of Kentucky.

I. INTRODUCTION

> When New England inventor Philip French had his epiphany 15 years ago, he didn't dream it would lead to an invention that would be pressed into service in a top-secret government project, or spawn an epic court battle over the limits of executive power. He was just admiring a tennis ball.[1]

The story behind the Federal Circuit's 2006 decision in *Crater Corporation v. Lucent Technologies* appears to be one of duplicity, culminating in the exploitation of French and two other inventors by the Navy and its contractor, Lucent.[2] At the behest of Lucent, the three inventors had shared their patented technology and other related information. After enthusiastically pronouncing their technology ideal for its naval project, however, Lucent denied compensation to the inventors. When they sought legal recourse, the Federal Circuit torpedoed their claims by applying the Military and State Secrets Privilege ("Privilege"). The decision created several serious problems, including a likely constitutional violation and a long-term threat to national security. Yet the *Crater* court failed to address either of these problems, which are increasingly likely to arise in light of the government's burgeoning use of the Privilege.[3] This Article examines these problems and suggests an alternative approach to applying the Privilege.

Few disagree with the need for the Privilege, which shields information that could endanger the security of the nation from public exposure.[4]

---

1. Kevin Poulsen, *Secrecy Power Sinks Patent Case*, WIRED, Sept. 20, 2005, http://www.wired.com/science/discoveries/news/2005/09/68894.

2. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260 (Fed. Cir. 2005), *cert. denied*, 547 U.S. 1218 (2006); Poulsen, *supra* note 1.

3. The Privilege has been invoked almost five times more often in the past eight years than during the height of the cold war. *Morning Edition*: *Administration Employing State Secrets Privilege at Quick Clip* (NPR radio broadcast Sept. 9, 2005), *available at* http://www.npr.org/templates/story/story.php?storyId=4838701 (citing Openthegovernment.org, which contended that from 1953-76, the Privilege was invoked only four times, but that from 2001-05, it was invoked twenty-three times); *see also* Robert M. Chesney, *State Secrets and the Limits of National Security Litigation*, 75 GEO. WASH. L. REV. 1249, 1291-92 (2007). Indeed, in the Judiciary Committee's report on the proposed State Secrets Protection Act, the Committee noted that "the Bush administration has raised the Privilege in over 25 percent more cases per year than previous administrations, and has sought dismissal in over 90 percent more cases." 154 CONG. REC. S198 (daily ed. Jan. 23, 2008) (statement of Sen. Kennedy).

4. Ellsberg v. Mitchell, 709 F.2d 51, 57 (D.C. Cir. 1983) (finding that "the state secrets doctrine pertains generally to *national security* concerns," and it is thus viewed

Unfortunately, the *Crater* court not only recognized the Privilege, but also relied on it to preclude the plaintiff from obtaining any meaningful discovery from Lucent and the Navy, effectively and unnecessarily shutting down the judicial process.[5] By permitting the Privilege to extinguish the plaintiff's claims, the *Crater* court triggered two concerns. First, rulings such as *Crater* prevent inventors from obtaining compensation when the government engages in unconstitutional "takings" of their trade secrets. This by itself should have given the *Crater* court pause. Second, the destruction of inventors' claims discourages military-related innovation.

Most significantly, the primary casualties of this approach are individual inventors and small companies, both of whom tend to be unfamiliar with the defense-industrial complex. This outcome runs counter to the Department of Defense's recently stated desire to expand procurement beyond the traditional large-company industrial base to increase access to cutting-edge communication and computer technology crucial to national security.[6] Yet, as demonstrated in *Crater*, powerful companies are the ones that benefit from a court's misapplication of the Privilege; this benefit stems from the knowledge that the Privilege gives them immunity from stealing individuals' and small business' innovations for military use. Thus, as one commentator noted, this case is "an especially pointed . . . lesson for [smaller] inventors who are contemplating approaching a company or others regarding licensing their technology," as well as a signal

---

"as both expansive and malleable"); *see* Jabara v. Kelley, 75 F.R.D. 475, 483 n.25 (E.D. Mich. 1977) (noting "the term 'military or state secrets' is amorphous in nature," but that it should be defined broadly, "referring to the military and naval establishments and the related activities of national preparedness"); Pollen v. Ford Instrument Co., 26 F. Supp. 583, 584 (E.D.N.Y. 1939) (noting the Privilege bolsters "the inherent right of self-preservation"); JOHN WIGMORE, WIGMORE ON EVIDENCE VIII 794 (John T. McNaughton rev., 1961) (noting the Privilege protects "matters relating to international relations, military affairs and public security").

5. Poulsen, *supra* note 1 (quoting Steven Aftergood, Director of the Federation of American Scientists' Project on Government Secrecy); *see supra* note 4.

6. OFFICE OF THE DEPUTY UNDER SEC'Y OF DEF. (INDUS. POLICY), TRANSFORMING THE DEFENSE INDUSTRIAL BASE: A ROADMAP (2003). The military has termed this focus on the integration of technology as "the Revolution in Military Affairs." *See e.g.*, Andrew F. Krepinevich, *Cavalry to Computer; the pattern of military revolutions*, THE NATIONAL INTEREST, Fall 1994, http://findarticles.com/p/articles/mi_m2751/is_n37/ai_16315042/; David Jablonsky, *U.S. Military Doctrine and the Revolution in Military Affairs*, PARAMETERS, Autumn 1994, *available at* http://www.carlisle.army.mil/usawc/parameters/1994/jablonsk.htm (early uses of the term for describing transformative change in military technology).

that long-term government goals can be taken hostage by short-term litigation strategy.[7]

This Article argues that an alternative strategy to the Federal Circuit's approach would avoid both the constitutional takings problem and the unnecessary encumbrance on the nation's defense goals. Rather than permitting the Privilege to result in an absolute concealment of information, courts should, where possible, apply more limited procedural safeguards that could protect sensitive information. In many situations, information could be divulged to the plaintiffs without wholesale public disclosure, thus permitting the case to proceed.

Part I of this Article describes how the *Crater* plaintiffs came to sue in federal court, and illustrates how more sophisticated commercial actors can take advantage of small inventors. This is accompanied by a discussion of how the development and evolution of the Privilege led to the *Crater* decision. In Part II, the Article considers the potential for a Takings Clause violation emanating from *Crater's* destruction of the value of the inventors' information, as well as the possibility of an analogous violation of patent rights. Part III explains that the damage is not limited to a purely legal injury; destroying the value of an inventor's efforts will discourage the technological innovation that the government has stated will be key to our nation's future defense plans. In light of the aforementioned problems, Part IV concludes that courts should, whenever possible, reject the government's attempt to suppress the designated information. Instead, courts should decide how to apply the Privilege only after taking into account both the nature of the privileged information involved and the effect of limited access to it. Finally, the Article outlines some less-restrictive protective measures that courts could apply. A bill currently under Senate consideration, the State Secrets Protection Act, proposes some of these measures.[8] However, this Article argues that those measures do not go far enough in allowing an inventor's claims to proceed in court. Courts should permit the dismissal of these claims only when no sufficiently protective procedural measures exist, because it is only in these circumstances that courts avoid both any latent takings problem and the unnecessary constraint on the incentive to invent.

---

7. Posting of Douglas Sorocco to Phosita: An Intellectual Property Law Blog, And You Thought You Were Having a Bad Day, http://dunlapcodding.com/phosita/2005/09/and-you-thought-you-were-having-a-bad-day.html (Sept. 22, 2005 21:51 EST) (commenting on the effective destruction of Crater's patent claims by the government).

8. State Secrets Protection Act, S. 2533, 110th Cong. (2008).

II. *CRATER CORPORATION V. LUCENT TECHNOLOGIES CORP.*

A. The Story of the Grab for an Individual Inventor's Technology

The story of Crater Corporation is a cautionary tale for inventors (particularly individual and small-business inventors) of devices with potential military applications. In 1991, Philip French and his co-inventors, Charles Monty and Steven Van Keuren, filed a patent application on a coupling device that links pipes or cables together without other complicated and expensive mechanisms.[9] They were having little success marketing their invention until they developed a commercial relationship with Lucent Technologies, a subsidiary of AT&T.[10] Lucent expressed an interest in obtaining a prototype of the mechanism so that it could evaluate it for use as an airtight "wetmate" in underwater fiber optic networks.[11] Lucent explained that, if appropriate, the device would be used to fulfill a classified contract with the United States.[12] Lucent sought both access to the inventors' undisclosed engineering drawings as well as a license to use the drawings to produce a prototype for research and development purposes.[13] According to the complaint, the inventors agreed to provide both the drawings and an R&D license; in return, Lucent agreed to keep the drawings secret, to negotiate another license agreement if they were interested in using the device in their network, and to provide the inventors with computer-assisted drawings of the technology.[14]

---

9. *See* U.S. Patent No. 5,286,129 (filed Aug. 7, 1991).

10. *See* Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW, 2007 WL 4593500, *2-3 (E.D. Mo. Dec. 28, 2007). The inventors originally garnered the interest of Plast-O-Matic, which was a subcontractor of Lucent on the fiber optic Naval project. *Id.* The inventors worked with Plast-O-Matic throughout 1994. *Id.* In early 1995, the inventors developed a direct connection to Lucent Technologies. Lucent Technologies, Inc. and AT&T Co. were both named as defendants in the lawsuit, although apparently all dealings occurred directly with Lucent. *Id.* Because there is no distinction in the Privilege analysis between the two defendants, the defendants are collectively referred to as "Lucent."

11. *Id.* at *4.

12. *Id.* ("Plaintiff also became aware that the project was under a United States government contract during the 'first third' or 'first quarter' of the relationship, and that this project was classified.").

13. *Id.*; Second Amended Complaint at ¶¶ 10–11, Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW (E.D. Mo. Aug. 31, 2006), 2006 WL 2699395.

14. Second Amended Complaint, *supra* note 13, at ¶¶ 12, 14–16. The district court held that there was no agreement to maintain the confidentiality of the technology. *Crater*, 2007 WL 4593500, at *3.

Shortly thereafter, Lucent rejected the inventors' request for CAD drawings of the coupler on the grounds that the inventors lacked the necessary security clearance required for access to the classified Navy project.[15] Nevertheless, the inventors were presumably excited to learn that Lucent's suitability testing indicated that the coupler was the best device for the project.[16] They were less excited, however, when Lucent offered only $100,000 to license the technology.[17] In May 1998, the inventors' corporation (Crater) filed a lawsuit alleging patent infringement, misappropriation of trade secrets, and breach of contract.[18]

In order to establish their patent infringement claim, Crater had to demonstrate that Lucent had fulfilled its government contract by manufacturing and using a device that fell within the scope of the patent.[19] By contrast, their trade secrets misappropriation claim required a demonstration of a disclosure of the inventors' confidential information to the government, even in the form of a non-infringing coupling.[20] To succeed on these claims, however, Crater would have had to either demonstrate that Lucent

---

15. *Crater*, 2007 WL 4593500, at *5.

16. Poulsen, *supra* note 1.

17. *Id.* One of the inventors wanted to demand $500,000 instead; the other, however, wanted to take the offered amount. *Id.*

18. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1261-62 (Fed. Cir. 2005), *cert. denied*, 547 U.S. 1218 (2006). Note that two of the inventors bought out the patent rights of the third inventor, and it is those two remaining patent holders who are the owners of Crater Corporation. Poulsen, *supra* note 1. The article uses "Crater Corporation" and "Crater" when referring to the party plaintiff, and "the inventors" when referring to their earlier inventive efforts. *Id.*

19. *See* 35 U.S.C. § 271 (2006); Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 535 U.S. 722, 730-32 (2002).

20. Missouri law defines misappropriation of a trade secret, in part, as:

> Disclosure or use of a trade secret of a person without express or implied consent by another person who:
> (a) Used improper means to acquire knowledge of the trade secret; or
> (b) Before a material change of position, knew or had reason to know that it was a trade secret and that knowledge of it had been acquired by accident or mistake; or
> (c) At the time of disclosure or use, knew or had reason to know that knowledge of the trade secret was:
> i. Derived from or through a person who had utilized improper means to acquire it;
> ii. Acquired under circumstances giving rise to a duty to maintain its secrecy or limit its use; or
> iii. Derived from or through a person who owed a duty to the person seeking relief to maintain its secrecy or limit its use.

MO. REV. STAT. § 417.453(2) (2008).

made some use of the inventors' patented or confidential information (the former for patent infringement, and the latter for trade secret misappropriation) to develop or produce the coupler, or demonstrate that Lucent shared this information with the government. Accordingly, the inventors sought discovery from Lucent and the government regarding their agreement and the nature of Lucent's coupler.[21]

Crater did not get far into litigation, however, when in August 1998, Lucent moved for dismissal of the patent infringement claim on the grounds that 28 U.S.C. § 1498(a) immunizes government contractors from patent infringement.[22] Section 1498 does not eliminate an inventor's ability to seek compensation for use of his patented inventions, however; instead it shifts liability directly to the government and expressly authorizes suits against the government for any such infringement.[23]

In September 2001, the Federal Circuit upheld the district court's dismissal of the patent claim.[24] As of 2007, Crater's § 1498(a) administrative

---

21. Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW, 2004 WL 3609347, at *2 (E.D. Mo. Feb. 19, 2004) (noting that the court entered a protective order sought by the government, "which stated that Crater could not 'conduct any discovery or serve any subpoena for information relating to the manufacture or use of plaintiff's coupling device, or any coupling device, by or on behalf of the United States' ").

22. Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW, 1999 WL 33973795 (E.D. Mo. Aug. 25, 1999). On appeal, the court stated that

> [i]n addition to giving the United States Court of Federal Claims exclusive jurisdiction over patent infringement suits against the government, § 1498(a) also provides "an affirmative defense for applicable government contractors." If a patented invention is used or manufactured for the government by a private party, that private party cannot be held liable for patent infringement.

Crater Corp. v. Lucent Techs., Inc., 255 F.3d 1361, 1364 (Fed. Cir. 2001) (citations omitted).

23. Section 1498 provides:

> Whenever an invention described in and covered by a patent . . . is used or manufactured by or for the United States without license of the owner thereof . . . the owner's remedy shall be by action against the United States in the United States Court of Federal Claims for the recovery of his reasonable and entire compensation for such use and manufacture.
>
> * * *
>
> For the purposes of this section, the use or manufacture of an invention described in and covered by a patent . . . by a contractor, a subcontractor, or any person, firm, or corporation for the Government and with the authorization or consent of the Government, shall be construed as use or manufacture for the United States.

28 U.S.C. § 1498(a) (2006).

24. *Crater*, 255 F.3d at 1363.

claim before the Department of Defense was still pending.[25] Since § 1498 only applies to patent claims, the trade secret and contract claims remained. Once the circuit court decided a disputed jurisdictional issue and the district court litigation resumed on the remaining claims, Crater tried to obtain previously-requested information regarding Lucent's coupler.[26] However, Crater found itself stymied on these claims as well. The government acknowledged that Lucent made 36 Crater-based prototypes for research and development, but claimed that these prototypes were rejected for use in the project.[27] Additionally, the government (having intervened in March 1999) asserted that neither it nor Lucent would produce information showing the selected design, because such information fell within the scope of the Privilege.[28]

The Privilege emanates from common law, most directly from a series of nineteenth century evidentiary rulings involving government concerns that "were woven together under the umbrella concept of a multifaceted 'public interest' privilege, some aspects of which were referred to under the subheading of 'state secrets.' "[29] The justification for the Privilege rests on the notion that the government possesses the right to shield information from the public when disclosure of such information would endanger national security (by permitting foreign entities to learn about the

25. Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW, 2007 WL 4593500, at *5 (E.D. Mo. Dec. 28, 2007); *see* 48 C.F.R. § 227.7004 (2008) (explaining the process of submitting an administrative claim for patent infringement with the Department of Defense).

26. *Crater*, 2004 WL 3609347, at *1. The district court further concluded that without federal question jurisdiction premised on the infringement claim, it lacked supplemental jurisdiction to hear Crater's state law claims for misappropriation of trade secrets and breach of contract, and therefore granted Lucent's motion to dismiss those claims as well. *Crater*, 1999 WL 33973795, at *1. On Crater's first appeal, the Federal Circuit affirmed the dismissal of the patent infringement claim, but reversed the dismissal of the state law claims. The Federal Circuit concluded that, under section 1367(c), the district court retained supplemental jurisdiction over the state law claims, and possessed the discretion as to whether or not to hear them. *Crater*, 255 F.3d at 1370-71. Thus, the Federal Circuit remanded to the district court for proceedings to determine whether to exercise such jurisdiction. *Id.* at 1371.

27. *Crater*, 2007 WL 4593500, at *5. Plaintiff contends that the Government had previously admitted making fifty-three prototypes. *Id.*

28. *Id.* at *1-2. Lucent did not take a position on whether the government properly invoked the Privilege. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1266 n.2 (Fed. Cir. 2005). Lucent argued that if the Privilege was upheld and the evidence regarding the contracted-for device was unavailable, the inventors' claims had to be dismissed. *Id.*

29. Chesney, *supra* note 3, at 1270-71.

country's military operations or diplomatic endeavors).[30] As one court described it, the Privilege is "the assertion of a paramount government right, the inherent right of self-preservation for purposes of national defense."[31] Under early United States common law, the Privilege was a generalized "public interest privilege" that covered communications between private informers and government officials (the informer's privilege) and intra-government communications (the deliberative process privilege), as well as communications related to national security and foreign relations (the military and state secret privilege).[32] As now applied, the Privilege is used to protect the country's defense functions, intelligence-gathering methods and abilities, and foreign diplomatic relations.[33]

While the Supreme Court formally recognized the Privilege in 1953,[34] the progenitor to the Privilege developed in the lead-up to, and midst of, the two world wars. In this period, weapon innovation became a more important facet of military power than ever before. Although technology has always acted as a "force multiplier," the ability of the military industrial

---

30. *See Crater*, 255 F.3d at 1370 ("The state secrets privilege allows the United States to 'block discovery in a lawsuit to any information that, if disclosed, would adversely affect national security.' ") (citing Ellsberg v. Mitchell, 709 F.2d 51, 56 (D.C. Cir. 1983)).

31. Pollen v. Ford Instrument Co., 26 F. Supp. 583, 584 (E.D.N.Y. 1939).

32. *See* Chesney, *supra* note 3, at 1271-80 (describing 19th century cases involved in the development of the Privilege, as well as early treatises discussing those cases). Professor Chesney has suggested that it may be possible to trace back the courts' generalized belief of the existence of the Privilege to the same origin—the English common law—as most privileges recognized by the courts, such as spousal privilege, attorney-client privilege, and the privilege against self-incrimination. *Id.* at 1273-74. As he notes, an 1836 treatise mentions two English cases that offer examples of an embryonic "matters of state" privilege. *Id.* at 1275-76.

33. *Crater*, 255 F.3d at 1370 (citing Ellsberg v. Mitchell, 709 F.2d 51, 56 (D.C. Cir. 1983)).

34. United States v. Reynolds, 345 U.S. 1, 12 (1953). In *Reynolds*, three employees of an Air Force contractor were killed when a B-29 Superfortress crashed. The employees' widows sued the government under the Federal Tort Claims Act. *Id.* at 3. During discovery, they sought production of accident reports concerning the crash, but were told by the Air Force that the release of such details would threaten national security. *Id.* at 3-4. Because of the failure of the government to produce the documents, a directed verdict in favor of the plaintiffs was granted by the trial court. *Id.* at 4. The United States Supreme Court reversed the decision (which had been affirmed by the Third Circuit), and remanded the case to the trial court. *Id.* at 12.

base to provide consistent innovation and progress in military technology became crucial to warfare in the twentieth century.[35]

In World War II, the ability of combatants to produce technologically advanced equipment and use it on the battlefield often meant the difference between victory and defeat.[36] It is no surprise, then, that the first 20th century case in which the Privilege was invoked (as well as two cases immediately prior to World War II) involved patent litigation by inventors whose devices had been allegedly used without authorization by the military or its contractors (armor-piercing projectiles in the first, apparatuses relating to gun-sighting in the later two).[37] While two of these cases involved claims brought against the government under § 1498 (or its predecessor), one case was a private patent suit in which the government intervened in order to invoke the Privilege, much like Crater.[38]

B. The Application of the Privilege in Crater

The government's strategy in district court was straightforward: assert the Privilege as to almost every document that the plaintiff sought to dis-

---

35. *See* RUPERT SMITH, THE UTILITY OF FORCE: THE ART OF WAR IN THE MODERN WORLD 132 (Knopf 2007) (elaborating the importance of technology to the mechanization of war on land, air, and sea).

36. *See* Mark Harrison, *The Economics of World War II: An Overview*, *in* THE ECONOMICS OF WORLD WAR II: SIX GREAT POWERS IN INTERNATIONAL COMPARISON 1, 26 (Mark Harrison ed., Cambridge Univ. Press 1998) (detailing and comparing the economic, industrial, and technological bases of the major combatants of World War II).

37. *See* Pollen v. United States, 85 Ct. Cl. 673 (1937); Pollen v. Ford Instrument Co., 26 F. Supp. 583 (E.D.N.Y. 1939); Firth Sterling Steel Co. v. Bethlehem Steel Co., 199 F. 353 (E.D. Pa. 1912).

38. *Compare Pollen*, 85 Ct. Cl. at 674 (action against the U.S. government) *with Pollen*, 26 F. Supp. at 584-85 (action against government contractor) and *Firth*, 199 F. at 353-54 (action against government contractor). Indeed, in some ways *Firth* seems similar to *Crater*. In this case, a company obtained a competitor's confidential technology directly from the Government. *Firth*, 199 F. at 353. Firth Sterling submitted a government bid for the manufacture of armor-piercing projectiles, which included non-public blueprints for its version of the product. *Id.* at 354. Nonetheless, the government then turned over Firth's blueprints to its competitor, defendant Bethlehem Steel, which used them to make their own version of the projectiles. *See id.* at 353. After Firth Sterling learned of this disclosure, it sued its competitor. *See id.* When Firth Sterling attempted to obtain discovery from both Bethlehem and the government, the Navy argued that "the researches and developments made by the Bureau of Ordnance and communicated to the Bethlehem Steel Company, for the purpose of fulfilling its contracts, embody secrets of military value to the government that could not be disclosed without detriment to the public interests." *Id.* at 354. Upholding the government's assertion of the Privilege, the court not only denied Firth's motion to compel the testimony, but impounded Firth's exhibits. *Id.* at 356.

cover. The government asserted the Privilege as to over 26,000 documents, including all information addressing whether or not any version or derivative of the Crater coupler was used.[39] To properly assert the Privilege, the head of the governmental department with authority over such information must formally invoke it and provide a declaration explaining the reason that national security would be at risk if the designated information were disclosed.[40] The government purportedly satisfied this requirement by submitting two declarations (one classified, one public) through the Secretary of the Navy.[41] The Navy's public declaration merely stated that the documents had been reviewed, and that the government had concluded that discovery proceedings could potentially provide classified information to U.S. adversaries. Furthermore, the public declaration stated that such disclosure could enable these adversaries to defeat U.S. military operations and gravely harm national security.[42]

The district court first attempted to address the Navy's concern by conducting an *in camera* inspection of the designated documents.[43] Given the vast number of documents, it is difficult to imagine that the court carefully considered each one's significance. Nonetheless, even a cursory re-

---

39. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1263-64 (Fed. Cir. 2005) (noting that the Government indicated that the concern was based on the possible disclosure of "information relating to the manufacture or use of [Crater's] coupling device, or any coupling device, by or on behalf of the United States").

40. *Id.* at 1265-66. In *Reynolds*, the Supreme Court had stated that invocation should only be made "after actual personal consideration by that officer." United States v. Reynolds, 345 U.S. 1, 8 (1953). But the *Crater* court concluded that this "personal consideration" did not demand that the department head personally review all of the information and documents sought. *Crater*, 423 F.3d at 1266. Although the *Crater* court did not elaborate, presumably the requisite consideration can be satisfied by the department head's personal consideration after being briefed by subordinates who reviewed the information and the security concerns. *See* Clift v. United States, 597 F.2d 826, 828-29 (2d Cir. 1979) (finding that the Privilege was sufficiently invoked by a mere Admiral's review of the documents, the Secretary of Defense Richard Chaney's declaration of review of the Admiral's affidavit and a sample of documents). The court noted, however, that "the Government would be wiser not to put courts to this test in the future." *Id.*

41. *Crater*, 423 F.3d at 1263-65. There were private and public versions of each of the declarations: one from Richard J. Danzig, then-Secretary of the Navy, and one from Acting Secretary of the Navy Hansford T. Johnson. *Id.*

42. *Id.* at 1263 (offering assertions by the government that compelling the government to turn over those documents "would permit potential adversaries to adopt specific measures to defeat or otherwise impair the effectiveness of those operations and programs," and "reasonably could be expected to cause extremely grave damage to the [country's] vital national security interests").

43. *Id.* at 1264.

view revealed that certain documents, such as publicly available court records, were not protected by the Privilege; as a result, the district court ordered a disclosure of such documents.[44] The government resisted disclosing even these documents, which prompted the court to order a "show cause" hearing regarding the use of the Privilege.[45] After the hearing, the court announced that even if the disputed documents were disclosed, the Privilege had indeed been properly invoked regarding specific documents that were crucial to proving the plaintiff's case.[46] As a result, not only did the district court prohibit the inventors from conducting any further discovery related to the manufacture or use of their invention, the court also concluded that:

> While Crater need not prove the government's intended use of the device, there is no way that Crater can prove misappropriation without showing that the defendants somehow incorporated Crater's design information into the classified government device. Because the protective order and the government's proper assertion of the state secrets privilege prevents Crater from discovering any evidence relating to this claim, Crater is unable to make out a prima facie case of misappropriation of trade secrets. Similarly, all of Crater's breach of contract claims except one require Crater to prove what the defendants did for the government.[47]

Lucent took no formal position as to propriety of the Privilege, but argued that its application would make it impossible to produce adequate evidence for a defense.[48] The district court agreed, declaring that even if the plaintiffs could somehow present a prima facie case, the government's invocation of the Privilege unfairly prevented Lucent from defending against Crater's claims.[49] As a result, the court concluded that the most just solution was to dismiss Crater's remaining misappropriation and contract claims.[50]

Considering the decision on appeal, the Federal Circuit began its analysis by acknowledging that the Supreme Court has previously indicated

---

44. *Id.* at 1265.
45. *Id.*
46. Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW, 2004 WL 3609347, at *2 (E.D. Mo. Feb. 19, 2004).
47. *Id.* at *3.
48. *Crater*, 423 F.3d at 166 n.2.
49. *Crater*, 2004 WL 3609347, at *4.
50. *Id.*

that the Privilege is "not to be lightly invoked."[51] Nonetheless, without even reviewing the documents themselves (relying instead on the lower court's decision, the government's declarations, and the parties' briefs), the Federal Circuit upheld the government's invocation of the Privilege for all of the documents.[52] Even more alarming was its intimation that, although such a remedy was "harsh," the district court was likely correct in dismissing the claims (because the claims could not prevail without the privileged information).[53] Nonetheless, the case was remanded for further consideration.[54] On remand, the district court found insufficient evidence to support Crater's claim that Lucent used the invention outside of the research and development prototypes and granted summary judgment to Lucent.[55]

---

51. *Crater*, 423 F.3d at 1265 (citing United States v. Reynolds, 345 U.S. 1, 7 (1953)).

52. *Id.* at 1269–70 ("None of us on this panel has inspected any of the information for which the claim is made."). *Crater* argued that the government's assertion of the Privilege was improper because the Secretary of the Navy and Acting Secretary of the Navy did not personally review the materials sought to be protected. *Id.* at 1265-66. The Federal Circuit rejected this argument, finding it sufficient that the Secretaries were informed of the nature and scope of the documents and determined, based on their personal knowledge, that disclosure would jeopardize a legitimate state secret and threaten national security. *Id.* at 1266; *cf.* Kinoy v. Mitchell, 67 F.R.D. 1, 9-10 (S.D.N.Y. 1975) (refusing to accept Attorney General's privilege claim in the absence of "an explicit representation" that he personally reviewed the material).

53. *Crater*, 423 F.3d at 1267 ("Although harsh, the presence of a properly invoked state secrets privilege requires dismissal of [a] claim that cannot prevail without the privileged information.") (citing McDonnell Douglas Corp. v. United States, 323 F.3d 1006, 1021 (Fed. Cir. 2003)).

54. Although the court recognized that "the record as it pertains to Crater's state law claims is not adequately developed," it then reviewed the claims (as indicated in the pleadings and briefs), and concluded that "as far as we can tell, it has not been established precisely what, if any, trade secrets exist in connection with the Crater coupler," or "whether, under Missouri law, there was a contract between Crater and Lucent and, if there was a contract, what its terms were." *Id.* at 1267-68.

55. Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW, 2007 WL 4593500, at *13 (E.D. Mo. Dec. 28, 2007). In opposition to Lucent's motion for summary judgment after this remand, Crater apparently changed its theory of recovery from a trade secret misappropriation claim to a "Submission of Ideas" claim available under New Jersey common law. *Id.* at *12. That claim required proving that: "(1) the idea was novel; (2) it was made in confidence, and (3) it was adopted and made use of" by the defendant. *Id.* Whereas establishment of a "trade secret" requires proof of other commercial use of the information by the owners prior to its alleged misappropriation, misuse of an "idea" does not. *Id.* On remand, it was undisputed that the Crater inventors had not sold or commercially used the coupler technology other than in regard to its dealings with Lucent. *Id.* at *3. "When trade secret claims do not meet the Restatement's 'use in busi-

While formally concurring in part and dissenting in part, Federal Circuit Court Judge Newman found several significant problems with the majority opinion. She began by disapproving of the decision to uphold the Privilege without reviewing the disputed information, in light of the Supreme Court's requirement that courts must determine for themselves whether the Privilege was appropriately claimed.[56] The failure to do so was particularly troubling to Judge Newman because she appeared to be skeptical that the government had used the Privilege appropriately, a skepticism fomented by several factors.[57] First, Judge Newman proclaimed that the government's declarations failed to sufficiently describe the nature of the documents and to identify the individual who reviewed these documents.[58] As a result, the court did not know whether the documents truly contained sensitive data. Further, the court was unable to ensure that the reviewer understood both the scope of the Privilege and the government's obligation to redact and separate sensitive information, producing the remaining non-sensitive information where possible.[59]

Second, Judge Newman observed that it was apparent that little or no effort was made to ensure that the Privilege was being used only on material that, if disclosed, would be a threat to national security. She noted that among the 26,000 purportedly "privileged" documents were items already in the public record.[60] This violates the tenet that the Privilege can only be used to protect information that threatens national security, and whenever possible, non-sensitive information must be separated out for

---

ness' requirement, they are treated as submission-of-idea cases." *Id.* at *12 (citing 49 NEW JERSEY PRACTICE, BUSINESS LAW DESKBOOK § 15:4 (Brent A. Olson et al. eds, 2007)). Nonetheless, when the Federal Circuit addressed the Privilege issue in 2006, it treated the inventors' information as potential trade secrets. Accordingly, this article analyzes the merits and problems of the circuit court's decision assuming that the inventors' information would indeed constitute trade secrets. *See infra* note 68.

56. *Cf. Reynolds*, 345 U.S. at 8; *Crater*, 423 F.3d at 1270 ("None of us on this panel has inspected any of the information for which the claim is made."); *see also* Molerio v. F.B.I., 749 F.2d 815, 822 (D.C. Cir. 1984) ("To some degree at least, the validity of the government's assertion must be judicially assessed.").

57. *Crater*, 423 F.3d at 1270.

58. *Id.*

59. *Id.* (noting the requirement that the official invoking the Privilege "must set forth, with enough particularity for the court to make an informed decision, the nature of the material withheld and of the threat to the national security should it be revealed") (citing Kinoy v. Mitchell, 67 F.R.D. 1, 8 (S.D.N.Y.1975)).

60. *Id.*

release.[61] Moreover, the Supreme Court in *Reynolds* stated that in instances where evidence is necessary, the Privilege should be scrutinized because executive officers cannot control judicial evidence.[62] Though the *Crater* majority gave lip service to this idea, it did not take its responsibility to impartial review seriously.[63] The privileged information was plainly necessary to the plaintiff in light of the district court's conclusion that the plaintiff's claims were unsustainable without it. In sum, the Crater majority failed to properly review the relevant documents to determine the applicability of the Privilege.

Regardless, as questionable as the specific applicability of the Privilege to the *Crater* documents might have been, this issue is outside of the scope of this Article. Of far greater concern are the larger constitutional and policy dangers, one of which was touched upon by Judge Newman and created by decisions such as *Crater*.

---

61. *Id.* (citing Ellsberg v. Mitchell, 709 F.2d 51, 57 (D.C. Cir. 1983)).

62. United States v. Reynolds, 345 U.S. 1, 9-10 (1953) ("Where there is a strong showing of necessity, the claim of privilege should not be lightly accepted."); *see* Molerio v. F.B.I., 749 F.2d 815, 822 (D.C. Cir. 1984) ("To some degree at least, the validity of the government's assertion must be judicially assessed."); Clift v. United States, 808 F. Supp. 101, 105 (D. Conn. 1991) (citing *Reynolds*, 345 U.S. at 9-10 and *Molerio*, 749 F.2d at 822). As early as 1953, Attorney General Herbert Brownell told President Eisenhower that classification procedures were " 'so broadly drawn and loosely administered as to make it possible for government officials to cover up their own mistakes and even their wrongdoing under the guise of protecting national security.' " Garry Wills, *Why the Government Can Legally Lie*, THE NEW YORK REVIEW OF BOOKS, Feb. 12, 2009, http://www.nybooks.com/articles/22285 (citing Brownell's June 15, 1953 letter to Eisenhower in KENNETH R. MEYER, WITH THE STROKE OF A PEN: EXECUTIVE ORDERS AND PRESIDENTIAL POWER 145 (Princeton University Press 2002)). And Erwin Griswold, President Nixon's former solicitor general, stated, "It quickly becomes apparent to any person who has considerable experience with classified material that there is massive overclassification, and that the principal concern of the classifiers is not with national security, but rather with government embarrassment of one sort or another." *Id.* (citing Erwin N. Griswold, Op-Ed., *Secrets Not Worth Keeping*, WASH. POST, Febr. 15, 1989, at A25).

63. It is unclear whether the Federal Circuit should have been applying a *de novo* or abuse of discretion standard. On other privilege issues, the Federal Circuit has deferred to the law of the regional circuit. *See In re* Pioneer Hi-Bred Int'l, Inc., 238 F.3d 1370, 1374 (Fed. Cir. 2001). But the Eighth Circuit has not clearly addressed the issue, and, regarding other privileges, "there is a conflict among courts of appeals on whether review of the district court's decision is *de novo* or for abuse of discretion." *Id.* (*comparing* Chaudhry v. Gallerizzo, 174 F.3d 394, 402 (4th Cir. 1999) ("We review the district court's decision that certain documents are subject to privilege *de novo* . . . "), *with In re* Grand Jury Proceedings, 219 F.3d 175, 182 (2d Cir. 2000) ("We review a district court's finding of waiver of the attorney-client and work-product privileges for abuse of discretion.")).

III. THE LEGAL AND POLICY IMPLICATIONS OF USING THE PRIVILEGE TO SUPPRESS INVENTORS' CLAIMS

The *Crater* majority's knee-jerk acquiescence to the government's invocation of the Privilege might have been expected. The military has invoked the Privilege in a variety of cases, such as soldiers' wrongful death claims.[64] Because a mere showing of a "reasonable danger" of security impairment will be successful, courts usually uphold the Privilege.[65] Moreover, because courts largely defer to the executive branch with regards to the preferred approach to safeguard security, they almost always suppress the designated information.[66] This is true in nearly all cases regarding unauthorized government use of intellectual property.[67]

---

64. *See, e.g.*, Schwartz v. Raytheon Co., 150 F. App'x. 627 (9th Cir. 2005) (referencing a *qui tam* whistleblower action filed under the False Claims Act, claiming that military contractor failed to satisfy requirements of a defense contract, in dismissing suit as a result of privilege); Zuckerbraun v. Gen. Dynamics Corp., 935 F.2d 544 (2d Cir. 1991) (dismissing an action brought against missile defense systems' manufacturers, designers, and testers for wrongful death of sailor who was killed when his ship was fired on by foreign aircraft after the invocation of the Privilege); *see also* Weston v. Lockheed Missiles & Space Co., 881 F.2d 814, 815 (9th Cir. 1989) (upholding the use of the Privilege in a suit by a homosexual employee of a government contractor against the contractor and the Department of Defense where the plaintiff claimed the government precluded the security clearance of the applicant revealing "evidence of homosexuality"); Darby v. U.S. Dep't of Defense, 74 F. App'x. 813 (9th Cir. 2003) (consisting of a suit regarding a denial of access to a military base where the government asserted the Privilege).

65. *Ellsberg*, 709 F.2d at 58 (citing *Reynolds*, 345 U.S. at 10).

66. *See, e.g.*, *Schwartz*, 150 F. App'x. at 628; *Weston*, 881 F.2d at 815; Hudson River Sloop Clearwater, Inc. v. Dep't of the Navy, No. CV-86-3292, 1989 WL 50794, at *4-5 (E.D.N.Y. May 4, 1989) (consisting of a challenge under the National Environmental Protection Act (NEPA) to the United States Navy's decision to construct a homeport in New York harbor which would allegedly include ships with nuclear weapons). The court upheld the Navy's invocation of the Privilege, suppressed the documents, and dismissed the suit. *Id.*

67. *See, e.g.*, *In re* Under Seal, 945 F.2d 1285, 1289 (4th Cir. 1991) (upholding a protective discovery order preventing the plaintiff from developing a factual basis for the claim of tortious interference with business, after a former government contractor sued several government agencies and a competing contractor, alleging that the defendant contractor and some government employees conspired to prevent the plaintiff's contract from being renewed); Clift v. United States, 597 F.2d 826, 829 (2d Cir. 1979); N.S.N. Int'l Indus. v. E.I. Dupont de Nemours & Co., 140 F.R.D. 275, 281 (S.D.N.Y. 1991) (consisting of a suit for a breach of contract regarding DuPont's use of NSN's technology for military armored personnel carriers in which the court upheld the military's claim of the Privilege solely based on the Secretary of Defense's declaration and suppressed sixty-six documents). The court claimed that the plaintiff was "in possession of substantial evidence from other sources supporting its claims." *Id. Contra* Halpern v. United States, 258

Courts' reluctance to consider alternative approaches to the Privilege may reflect, positively, confidence in the executive's national security expertise and, negatively, an abdication of decision-making when consequences could be potentially disastrous. Whatever the reason, the outcome in *Crater* (the suppression of all of the designated documents and dismissal of the claims) creates two substantial problems.[68] First, the uncompensated use of an inventor's trade secrets is likely an unconstitutional violation of the Takings Clause. Second, the abuse of inventors' efforts will deter small businesses from engaging in the development of military applications, which derails a separate defense policy goal. Both problems demonstrate why national security concerns should not diminish the judiciary's role in checks and balances.

F.2d 36, 44 (2d Cir. 1958) (holding that "the privilege . . . is inapplicable when disclosure to court personnel in an in camera proceeding will not make the information public or endanger the national security"). For instance, in *Clift v. United States*, the plaintiff sought compensation for the government's alleged use of his patented cryptography invention (as well as compensation for a claim under the Invention Secrecy Act, which is required when the government temporarily or permanently places a patent application under seal). 808 F. Supp. 101, 102-03 (D. Conn. 1991). The Second Circuit upheld the government's assertion of the Privilege and suppressed the privileged information, which resulted in the eventual dismissal of the plaintiff's claims. *Clift*, 597 F.2d at 829. In vacating the order of dismissal, the Second Circuit stated that Clift could perhaps proceed with his case without the requested documents or that, at some unforeseen point in time, the disclosure of the requested documents would no longer imperil the national security. *Id.* at 830. Moreover, the circuit court encouraged the government to "be as forthcoming as it can be without risk to the national interest." *Id.* Yet the district court later dismissed the claims. *Clift*, 808 F. Supp. at 111; *see also* Pollen v. United States, 85 Ct. Cl. 673, 674 (1937) (upholding the Privilege to preclude the plaintiffs from subpoenaing the testimony of several witnesses).

68. As explained above, on subsequent remand, Crater changed its theory of recovery to a "submission of idea" claim. *See supra* note 55. Whereas establishment of a "trade secret" requires proof of other commercial use of the information by the owners, an "idea" does not. Crater Corp. v. Lucent Techs., Inc., No. 4:98CV00913 ERW, 2007 WL 4593500, at *12 (E.D. Mo. Dec. 28, 2007). However, when the Federal Circuit addressed the Privilege, it treated the inventors' information as trade secrets. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1262 (Fed. Cir. 2005). While the circuit court expressed skepticism of the inventors' ability to prove their claims, there is no indication that its privilege analysis was based on a clear rejection of the possibility that plaintiff might have possessed a trade secret. Moreover, given the likelihood that some unauthorized government use of information will likely run afoul of undisputed trade secrets, the Federal Circuit's application of the Privilege continues to be of concern. Accordingly, this Article analyzes the merit and problems of the circuit court's decision assuming that the inventors' information would indeed have constituted trade secrets.

A. The Likelihood of a Takings Clause Violation

The government's use of trade secrets should immediately raise a red flag as possibly violating the Takings Clause, which prohibits the government from seizing of or trespassing on private property without compensation.[69] Indisputably, it is a longstanding Supreme Court doctrine that many property interests do not qualify as protected "property" under the Takings Clause.[70] Here, however, the right is clear: in 1984, the Supreme Court held in *Ruckelshaus v. Monsanto* that trade secrets are indeed "property" for purposes of the Takings Clause, and that the government's "trespasses" on an owner's trade secrets can constitute a taking.[71] In *Crater*, if the complaint accurately describes Lucent's agreement to limit its use of the plaintiff's information to its testing of the coupler, and if that information did constitute protectable trade secrets, any further use of the information to produce a coupler on behalf of the government would likely establish a taking.[72] Yet, because the application of the Privilege effectively

69. For almost 150 years, takings claims were limited to *physical* seizure or trespass. *See* Davida H. Isaacs, *Not All Property Is Created Equal: Why Modern Courts Resist Applying the Takings Clause to Patents, and Why They Are Right To Do So*, 15 GEO. MASON. L. REV. 1, 17 n.89, 29-30 (2007) (noting Justice Blackmun and constitutional scholars who have described the nineteenth century understanding of the Takings Clause). But in 1922 the doctrine of "regulatory takings" was established when the Supreme Court declared that the Takings Clause could encompass losses to property value resulting from other government action. *See* Lingle v. Chevron U.S.A. Inc., 544 U.S. 528, 537 (2005) (referring to *Mahon*); *see also* Pa. Coal Co. v. Mahon, 260 U.S. 393 (1922) (holding that a diminution in property value could be so great that it was "the functional equivalent of a 'practical ouster of [the owner's] possession' ").

70. *See* Isaacs, *supra* note 69, at 36 ("[b]eing property is a necessary requirement for Takings Clause protection, but it is not a sufficient one"). Isaacs identifies another scholar who has commented that, "constitutional scholars know that merely classifying a legal entitlement as property is insufficient by itself to justify providing its owner with the full panoply of constitutional remedies." *Id.*

71. Ruckelshaus v. Monsanto Co., 467 U.S. 986, 1001 (1984).

72. Under Missouri's law, in order for the government to have committed a violation, the government also would need to have reason to know that the information was misappropriated. MO. REV. STAT. § 417.453(2) (2008). If the trade secrets might still have some commercial value, the government's unauthorized use would not constitute a clearly compensable "total taking." Unfortunately, the Court has been able to clearly articulate neither what qualifies as deprivation of "*all* economically beneficial us[e] of [the] property" sufficient for a total, per se taking, nor how to judge when partial economic diminution is sufficient to constitute a taking. *Lingle*, 544 U.S. at 538 (emphasis in original). "Indeed, shortly before retiring, Justice O'Connor pulled no punches: 'Our regulatory takings jurisprudence cannot be characterized as unified.' " Isaacs, *supra* note 69, at 26 n.143 (citing *Lingle*, 544 U.S. at 539; Palazzolo v. Rhode Island, 533 U.S. 606, 631 (2001); Tahoe-Sierra Pres. Council, Inc. v. Tahoe Reg'l Planning Agency, 535 U.S. 302,

eliminates the trade secret owner's claim, the inventors were prevented from receiving the compensation to which they were constitutionally entitled.

There is one possible argument against a trade secret owner's takings claim. One of the Court's few clear declarations regarding takings law has been that the government need not compensate even the complete destruction of property value if the justification comports with certain "background principles."[73] The theory is that the property owner should recognize that such background principles have the potential to diminish the value of his property, and thus any determination of property value without consideration of those principles is not reasonable.[74] In the real property context, this term refers to traditional property and nuisance common law.[75] In *Ruckelshaus*, the Court concluded that these background principles include not only the general laws applicable to that particular type of property, but also to the regulatory environment of the trade secrets involved.[76] Assuming that a court would consider the protection of national security to be a background principle (after all, a fundamental purpose of the national government is to protect national security), the government's use might not constitute a taking. However, the background principles rationale has been applied only in the context of regulatory takings (takings in which government legislation or regulation indirectly decreases the value of the property), and not in the context of direct takings (takings where the government actually uses the property).[77] It would significantly change

---

330-31 (2002)). In any case, the evidence indicates that the inventors had not been able to create other commercial interest in the coupler technology; thus, the trade secrets' only apparent value involved their Navy application.

73. *See Lingle*, 544 U.S. at 538 (stating that in cases of such "categorical" takings, "the government must pay just compensation for such 'total regulatory takings,' except to the extent that 'background principles of nuisance and property law' independently restrict the owner's intended use of the property") (citing Lucas v. S.C. Coastal Council, 505 U.S. 1003, 1026-32 (1992)).

74. *Lucas*, 505 U.S. at 1029.

75. *Id.*

76. *Ruckelshaus*, 467 U.S. at 1006 (noting the regulatory landscape that put Monsanto on notice that the EPA could reveal Monsanto's formulae of products, as well as its health, safety, and efficacy data, to Federal agency and the public).

77. See Nestor M. Davidson, *Standardization and Pluralism in Property Law*, 61 VAND. L. REV. 1597, 1659 (2008) (noting the courts' use of the "background principles" in cabining regulatory takings); *see also* BRIAN W. BLAESSER & ALAN C. WEINSTEIN, FEDERAL LAND USE LAW & LITIGATION § 3:38 (updated 2009) (noting background prin-

takings doctrine if the government were permitted to seize property and assert national security as a basis for defending against a resulting takings claim.

This brings us to the practical "catch-22": even if the trade secret owner were to pursue a viable takings claim, the government could presumably shut down any such claim by invoking the Privilege. This is consistent with the policy behind the Privilege; if national security would be threatened by the disclosure of information during a trade secret suit, the same threat would exist if disclosure occurred during a takings suit.[78]

One might suggest that this menace is merely phantom because inventors usually protect their inventions through patent rights instead of trade secret rights. If inventors are unable to receive compensation through the government's administrative process (a meaningful concern for inventors), then they are left with the option of asserting a § 1498 compensation claim.[79] But the same problem arises: there is no reason that the govern-

---

ciples as the exception to the "the categorical rule that total economic deprivation by regulation is a taking").

78. That raises a further question: Could the invocation itself constitute a taking? Research has uncovered no other use of the privilege that resulted in a takings claim. It remains unresolved whether any judicial ruling can provide the basis for a takings claim. *See* J. Nicholas Bunch, Note, *Takings, Judicial Takings, and Patent Law*, 83 TEX. L. REV. 1747, 1752 n.29, 1760-63 (2005) (noting that while the concept of a judicial taking finds support in *Pruneyard Shopping Center v. Robins*, 447 U.S. 74, 82 (1980), a concurrence by Justice Stewart in *Hughes v. Washington*, 389 U.S. 290, 296-97 (1967), and some language in *Hartford-Empire Co. v. United States*, 323 U.S. 386, 415 (1945), which suggested that a court order decreeing the forfeiture of a patent as a remedy for antitrust violations would violate the Takings Clause, "the fact remains that the Supreme Court has yet to find a case in which a judicial taking has occurred"). Bunch argues that Federal Circuit decisions that dramatically depart from settled precedent should give rise to judicial takings claims. *Id.* at 1754-55. However, there is the reasonable concern that if any action by the judicial branch of government could instigate a constitutional taking, then every court loss would leave the government open to liability. But the military and state secrets privilege does appear to be an anomaly that could distinguish it from other judicial determinations—the courts have held that once where national security is at risk, acceptance of the Privilege is not discretionary. *See, e.g.*, United States v. Reynolds, 345 U.S. 1, 10 (1953); Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1266 (Fed. Cir. 2005). Thus, there is a strong argument that the destruction of the claim, though applied by the judiciary, is primarily an executive branch activity.

79. As indicated by the apocryphal discussion offered by one patent attorney, the government is often not forthcoming with compensation to small businesses. *See* Posting of Michael L. Slonecker to Patently-O, Patently-O TidBits, http://www.patentlyo.com/patent/2007/03/patentlyo_tidbi_2.html#comment-64529722 (Mar. 27, 2007 15:23:00 CST) (comment); discussion *infra* at Section III.B. Lucent almost certainly knew that it

ment would invoke the Privilege in a private patent litigation and not invoke it in a § 1498 case involving the same technology (this might occur *if* the security risk was fabricated to protect the contractor from liability for political reasons).[80] Finally, the last recourse, a Takings Clause claim, would likely run into the same problem as described above: the government's invocation of the Privilege would effectively preclude a patent holder from recovering any damages.[81]

---

would be immunized from liability by § 1498 for any use that it made of the coupler under its government contract. Therefore, it had little motivation to avoid infringement.

80. The Court of Federal Claims is a public forum just like any district or circuit court. *See* Pollen v. United States, 85 Ct. Cl. 673, 678 (1937) (noting that the jurisdiction of the court was predicated on the 1910 statutory predecessor to § 1498). In *Pollen*, the plaintiffs sought recompense for the Navy's appropriation of two devices, the Federal Court of Claims upheld the privilege to preclude plaintiffs from subpoenaing the testimony of several witnesses—one a government witness and one a private witness who obtained information through a "confidential relationship" with the military. *Id.* at 680. At the time, the Judicial Code permitted "[the] head of any department [to] refuse and omit to comply with any [subpoena from the court] for information or papers, when, in his opinion, such compliance would be injurious to the public interest." *Id.* at 677. The court rejected the argument that the Government waived its right to invoke a privilege which would effectively vitiate the claim, concluding that the section provides only an opportunity to sue the government, using the same evidentiary rules as in any other court. *Id*. The court also rejected the assertion that application of the Privilege would violate plaintiffs' due process rights. *Id.* at 676. It wasn't until far after *Pollen*—to be exact, in the 1999 Supreme Court's *Florida Prepaid* decision—that it was even clear that a patent holder could, in fact, assert due process claims.

81. In any case, getting past the Privilege hurdle would hardly be a guarantee of success of such a claim, because the Federal Circuit recently held that patents are not "property" for purposes of the Takings Clause. Zoltek Corp. v. United States, 442 F.3d 1345, 1353 (Fed. Cir. 2006) (per curiam), *cert. denied*, 127 S. Ct. 2936 (2007). That ruling was highly controversial, with a substantial number of practitioners and scholars strongly arguing that the earlier Supreme Court statements mandated a contrary result. There were some expectations that either the Federal Circuit would reverse this panel decision en banc, or that the Supreme Court would accept certiorari and reverse the lower court. Neither happened. Isaacs, *supra* note 69, at 13 ("[A] gambler might have reasonably wagered that the Federal Circuit would grant a rehearing en banc and overrule the panel."). *Compare* Isaacs, *supra* note 69, at 43 (noting that if patents are considered federal benefits, they should not be protected by the Takings Clause), *with* Adam Mossoff, *Patents as Constitutional Private Property: The Historical Protection of Patents Under the Takings Clause*, 87 B.U. L. REV. 689, 690 (2007) (arguing that patents are Takings Clause property); Shubha Ghosh, *Reconciling Property Rights and States' Rights in the Information Age: Federalism, the "Sovereign's Prerogative" and Takings after* College Savings, 31 U. TOL. L. REV. 17, 41 n.163 (1999) (arguing that patents are property for purposes of the Takings Clause).

In sum, by permitting the government's application of the Privilege to eradicate the plaintiff's claims, the court wrongly prevented the plaintiff from obtaining constitutionally mandated redress for the government's use of the plaintiff's trade secrets. Moreover, even a direct claim under the Takings Clause would presumably be thwarted by the Privilege. Thus, as applied in *Crater*, the Privilege would be more powerful than constitutional protections. And yet while one might be able to engage in a debate as to if (and when) national security should trump constitutional guarantees, the court failed even to consider this imbalance. If the *Crater* majority recognized these conflicting obligations, they might have considered alternatives, such as those discussed in Part III, to the drastic result of denying inventors the ability to obtain relief.

B. Undermining Small Businesses' Incentive to Innovate Runs Counter to the Department of Defense's National Defense Strategy

As explained above, the *Crater* court's application of the Privilege created a highly troubling constitutional problem. Given Judge Newman's previous vehement support for an inventor's right to Takings Clause protection, one might have expected that she would have directed her attention in *Crater* to the potential Takings Clause violation.[82] Yet Judge Newman focused on another effect of the Privilege: the danger of discouraging military innovation.

While acknowledging the importance of protecting military and state secrets from the nation's enemies, Judge Newman declared that, "[a]t the same time, persons who serve the government must have a reasonable way of resolving disputes."[83] Providing the *Crater* inventors with legal recourse should not have been merely altruistic; rather, the fair resolution of disputes is necessary to ensure the government's continued access to the private sector's talents.[84] Judge Newman recognized that patent and trade secret protections exist to provide incentives to potential inventors. With-

82. Upon the Federal Circuit's refusal to re-hear *Zoltek* en banc, it was Judge Newman who offered a strong dissent criticizing the Federal Circuit for failing to protect patentholders' property rights. *See* Isaacs, *supra* note 69, at 14 (noting that in her dissent to the denial to rehear *Zoltek* en banc, Judge Newman argued that recognition of the right to a Takings Clause remedy for trade secrets while rejecting a similar remedy for patents was in conflict with Supreme Court and lower court precedent that referred to patents as "property").

83. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1270 (Fed. Cir. 2005), *cert. denied*, 547 U.S. 1218 (2006).

84. *Id.* at 1271.

out such protections, others would simply copy an invention and undersell the original inventor. The right to temporarily charge monopoly prices encourages people to invest more money and effort into innovation. If government assertions of the Privilege are permitted to destroy such a monopoly's value, inventors will lose this incentive.

Society may be willing to forego the creation of some relatively inconsequential inventions, but as Judge Newman indicated, it cannot afford to intentionally remove the intellectual property incentives for inventions with military application (such as the coupler).[85] Judge Newman is not the first jurist to recognize the danger created by precluding inventors from obtaining recompense for their inventions. Even while upholding the government's assertion of the Privilege in *Clift*, the Second Circuit Court of Appeals noted that the destruction of inventors' claims conflicts with the policy goal of encouraging inventors to pursue inventions with military potential.[86]

Most problematically, the disincentive caused by the dismissal of such claims falls disproportionately on smaller inventors. Permitting the Privilege to thwart the small inventors' incentive to innovate is particularly disturbing because the Department of Defense has recently recognized that such inventors are playing an increasingly important role in national security. Their significance is apparent in the "Revolution in Military Affairs" report, which represents the integration of advanced communications, computer, and information technology into military hardware and military institutions.[87] Much of this newer technology lies outside the expertise of traditional members of the defense-industrial complex.[88] Consequently, the Defense Department has suggested that it expects to rely more on non-traditional providers (which tend to include smaller inventive entities) than

---

85. *Id.*

86. Clift v. United States, 597 F.2d 826, 829 (2d Cir. 1979) (citing Halpern v. United States, 258 F.2d 36, 39 (2d Cir. 1958)). In vacating the order of dismissal, the Second Circuit stated that Clift could perhaps proceed with his case without the requested documents or that, at some unforeseen point in time, the disclosure of the requested documents would no longer imperil the national security. *Id.* at 830. Moreover, the circuit court encouraged the government to "be as forthcoming as it can be without risk to the national interest." *Id.*

87. Joint Chiefs of Staff, Joint Vision 2010 11-16 (1996); *see also* William S. Cohen, Dep't of Def., Annual Report to the President and Congress, ch. 10 (1999), *available at* http://www.dod.mil/execsec/adr1999/chap10.html.

88. *See* Peter Dombrowski & Eugene Gholz, Buying Military Transformation: Technological Innovation and the Defense Industry xi (2006).

it has in the past.[89] This expectation of a shift in innovation's origin is the result of Pentagon studies conducted in the early 2000s. These studies emphasized the need to retain technological superiority in numerous areas in order to maintain "warfighting leadership."[90] They specifically called for the United States to expand access to innovative technologies available in the commercial realm, as this access would increasingly provide capability and cost advantages to the Defense Department.[91] Moreover, these studies concluded that gaining the benefit of private sector innovation depends on ensuring competition from firms that are not part of the traditional defense industrial base.[92] One study identified twenty-four small or non-traditional firms that could be expected to contribute to maintaining U.S. technological advantages.[93] As a result, the Department of Defense has made the transformation of the defense-industrial base by reaching out to smaller inventive entities a priority.[94]

There is no question that the government should be encouraging small businesses to invest more in innovation with potential military applications. However, as some scholars have noted, traditional members of the defense-industrial complex have substantial advantages over newcomers.[95] Larger, well-established companies are regularly involved in producing numerous goods and services used by the government.[96] These companies benefit from their familiarity with the remarkably time consuming, complex landscape of defense procurement. Indeed, former Secretary of Defense Donald Rumsfeld noted that the delay caused by the process of defense procurement benefited larger companies because they have access to lawyers and lobbyists that smaller companies do not.[97] Moreover, gov-

---

89. Off. of the Deputy Under Sec'y of Def. (Indus. Policy), Transforming the Defense Industrial Base: A Roadmap 8-13 (2003) [hereinafter Roadmap].

90. Off. of the Deputy Sec'y of Def. (Indus. Policy), Defense Industrial Base Capabilities Study: Command and Control vi (2004).

91. *Id.* at ix.

92. *Id.* at 21; ROADMAP, *supra* note 89, at 35.

93. ROADMAP, *supra* note 89, at v.

94. Donald Rumsfeld & Rik Kirkland, *Don Rumsfeld Talks Guns and Butter; Setting Priorities When Stakes are Scary*, FORTUNE, Nov. 18, 2002, at 143. Rumsfeld noted, critically, that "[the] government tends not to have the kind of interaction with [all of] the creativity and innovation that exists in our society." *Id.*

95. *See, e.g.*, DOMBROWSKI, *supra* note 88, at 24-25 (explaining defense firm hiring practices).

96. *See, e.g.*, *Id.* at 24 (explaining the large firm advantage).

97. Rumsfeld & Kirkland, *supra* note 94 ("Delay helps the big companies, because they've got all the lawyers and all the lobbyists and all the people in Washington. Smaller companies don't have time to do all of that.").

ernment officials are generally interested in maintaining good relations with these regular suppliers. These companies further ensure favorable treatment by employing former Department of Defense and military personnel.[98] As a result, even when one of these large companies holds the intellectual property rights to an invention useful for a military system and is *not* the party contracted to produce that system, the military is likely to agree to license that invention rather than try to undercut one of its regular suppliers.

By contrast, smaller businesses do not have the political and economic influence to compel the government to license their intellectual property. Instead, they may be left feeling impotent and apprehensive. Indeed, one commenter on the well-respected Patently-O blog described this impotence and apprehension through the following apocryphal dialogue:

> Patentee sends letter to [the government] contractor saying "You are infringing my patent." [sic]
>
> Contractor sends letter to patentee saying "Pound sand, weasel. Go take it up with the [the government,] 'cuz it ain't my problem."
>
> Patentee send letter to [the government] saying "My patent is being infringed by the work being done under your contract with the contractor."
>
> [The government] replies "So sue me. I will consign your claim to a 'black hole' where it will languish for years as an administrative claim."
>
> Patentee send letter saying "Here is my claim."
>
> [The government] replies "Got it. We will get back to you sometime in the next decade of [sic] so. Hopefully, by then you will be under Chapter 11."[99]

To the extent that this view is prevalent among small businesses, they will anticipate the need to rely on legal recourse rather than clout. As a

---

98. DOMBROWSKI, *supra* note 88, at 24.

99. Posting of Michael L. Slonecker to Patently-O, Patently-O TidBits, http://www.patentlyo.com/patent/2007/03/patentlyo_tidbi_2.html#comment-64529722 (Mar. 27, 2007 15:23:00 CST) (comment).

result, the incentive for smaller companies to develop technology beneficial to the Department of Defense will depend on robust intellectual property rights.

The government wants small inventors of newer technology to focus their innovative efforts with defense applications in mind, and holds out the promise of substantial government sales as motivation. Innovators attentive to that prospect could be dissuaded by the possibility that, if the fruit of their efforts are utilized without authorization, their avenues of recourse could be blocked through the government's use of the Privilege. Therein lies the problem: the government is anticipating an increased reliance on small entity innovation to provide key components of many forthcoming national defense systems, yet its use of the Privilege to foster uncertainty in such entities is dangerously counterproductive.

In sum, as Judge Newman argued: "[f]air resolution of disputes is necessary to ensure the government's continued access to the private sector's talents."[100] Unexpected termination of disputes through the Privilege undermines such a fair resolution. However, it is even more counterproductive than perhaps even Judge Newman realized, because the government is expecting to rely more heavily on individual and smaller entities for much of its new defense technology. These groups lack the political clout and long-term institutional relationships to ensure that the military will properly license their technology. As a result, unless they can rely on having their intellectual property protected, small inventors may be disinclined to pursue inventions sought after by the military, or to make their non-patented civilian inventions available to the military. Thus, injudicious application of the Privilege discourages innovation by those that the defense department most wishes to encourage.

IV. RESOLVING THE COMPETING NEEDS TO ENSURE NATIONAL SECURITY AND ENCOURAGE INVENTORS

One might argue that the discouragement of inventors, especially smaller inventors, is an unfortunate but inevitable disadvantage of the use of the Privilege. But it is not always inevitable, and it would be preferable to avoid that consequence to the greatest extent possible. While it is true that *Reynolds* suggests that the need to protect military secrets trumps the evidentiary needs of some plaintiffs, this exhortation does not require that the only remedy that a court may consider is a complete bar to disclosure

100. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1270 (Fed. Cir. 2005).

of the information.[101] Thus, faced with possible constitutional violations and long-term security concerns, it is within the discretion of, and incumbent upon, Congress and the courts to fashion an alternative response to the invocation of the Privilege.[102] This approach, briefly mentioned by Judge Newman in her partial dissent, deserves further consideration.[103]

Similar executive use (and arguably abuse) of the Privilege in cases regarding terrorism and national security investigations[104] has recently prompted a bipartisan group of Senators to propose a bill seeking to provide guidance to federal courts considering the assertion of the state secrets privilege.[105] This legislation, the State Secrets Protection Act, would provide increased judicial oversight of the government's use of the Privilege.[106]

The majority report emphasizes that this bill "codifies the principle that state secrets assertions are justifiable" and that ''[t]he court itself must determine whether the circumstances are appropriate for the claim of pri-

---

101. United States v. Reynolds, 345 U.S. 1, 11 n.26 (1953) (citing Totten v. United States, 92 U.S. 105 (1875)).

102. *Crater*, 423 F.3d at 1271 ("The judicial obligation is to enable resolution, with safeguards appropriate to the subject matter."); Clift v. United States, 808 F. Supp. 101, 107 (D. Conn. 1991) ("The effect that a successful invocation of the privilege has on a case varies; the court must invariably 'consider whether and how the case may proceed in light of the privilege.' ") (citing Fitzgerald v. Penthouse Int'l, Ltd., 776 F.2d 1236, 1243 (4th Cir. 1985). However, the *Clift* court erroneously presumed that invocation of the privilege meant that, at a minimum, the court was forced to remove the sensitive information from the case completely. *See id.*

103. *Crater*, 423 F.3d at 1271 ("We should remand this case for *in camera* proceedings that would protect the information from public disclosure, and allow this dispute to come to closure. Trials *in camera* of issues subject to secrecy restraints are not new, and such trial would be the appropriate procedure in this case."); *see also* Pollen v. United States, 85 Ct. Cl. 673, 680 (1937) ("[T]he plaintiffs' motion for the court to instruct the defendant to proceed with the case *in camera* must be overruled.").

104. As the majority report makes clear, these cases primarily involve either government prosecution or private suits seeking the release of information about the government's investigative activities. S. REP. NO. 110-442, at 3 (2008).

105. Senate Judiciary Committee Chairman Patrick Leahy (D-Vt.), and Members Arlen Specter (R-Pa.), Russ Feingold (D-Wis.), Edward Kennedy (D-Ma.), Sheldon Whitehouse (D-R.I.), and Claire McCaskill (D-Mo.) introduced the bill on February 11, 2009. Press Release, U.S. Senator Patrick Leahy, Leahy, Specter, Feingold, Kennedy Introduce State Secrets Legislation (Feb. 11, 2009), *available at* http://leahy.senate.gov/press/200902/021109b.html. A version of the bill was previously introduced in January 2008, but died in that Congressional session after approval by the Judiciary Committee. State Secrets Protection Act, S. 2533, 110th Cong. (2008).

106. State Secrets Protection Act, S. 417, 111th Cong. (2009).

vilege."[107] For example, the bill requires judges to review the allegedly privileged evidence before dismissing a case, instead of relying on the government's assertions as to the nature of the proposed privileged information.[108]

The bill also formalizes procedures for invocation of the Privilege, including requiring government affidavits offering personal knowledge of the basis for the Privilege.[109] All purportedly privileged documents, along with an index, must be made available to the court.[110] Moreover, by requiring that most hearings based on the assertion of the Privilege be conducted in camera,[111] and if necessary, by cleared special masters,[112] the bill undercuts the government argument that sharing the information needed to analyze the Privilege with the court would, in and of itself, impair national security because of the public's access to a district court's records or the judge's own lack of security clearance.[113]

Furthermore, the bill discourages attempts by the government to seek ex parte hearings by instructing the government to provide, where possible, opposing parties' attorneys with the needed national security clearances;[114] where such clearances are not possible (presumably because of

---

107. S. REP. NO. 110-442, at 24-25 (citing United States v. Reynolds, 345 U.S. 1, 8 (1953)). The majority emphasized that "[t]o the extent that *Reynolds*"—in which the court eventually rejected the opportunity to review the purported privileged—"has been read to bless the judicial practice of not 'insisting upon an examination of the evidence, even by the judge alone, in chambers' under certain conditions, the bill overrides the court's discretion to adopt such a practice." *Id.* at 24 (quoting *Reynolds*, 345 U.S. at 10).

108. *Id.* at 20-22. A court may review ''a sufficient sampling of the [allegedly privileged] evidence'' if such a review will permit proper scrutiny of the documents at issue. *Id.* at 22.

109. *Id.* at 12. The majority report notes that:

> The Government may assert the privilege to withhold information in discovery or to prevent the disclosure of information in litigation. The assertion is made through an affidavit signed by the head of the relevant agency and submitted to the court. An unclassified version of the affidavit must be made public.

Id.

110. *Id.* at 22.

111. *Id.* at 18. The proposed statute permits a public hearing if the only issues to be addressed are purely legal and do not necessitate the disclosure of sensitive factual information. *Id.*

112. *Id.* at 19-20.

113. *Cf. infra* notes 129-130

114. S. REP. NO. 110-442, at 19-20. *But see* Halkin v. Helms, 598 F.2d 1, 7 (D.C. Cir. 1978) (rejecting the possibility that plaintiffs' counsel be present when the privileged information, interceptions of plaintiffs' foreign communications by the National Security

the attorneys' background), the court is authorized to appoint a guardian ad litem to represent the private party's interests.[115] Finally, when material evidence is found to be privileged, the legislation compels a court to order the government to create, if possible, a non-privileged substitute for the evidence, such as an unclassified summary or a redacted version; government refusal to do so would oblige a court to decide the relevant issue of fact or law against the government.[116] Some combination of these procedural approaches has been used in prior individual cases. However, as demonstrated by the *Crater* courts' failure to review the alleged privileged information, these "best practices" have not been consistently implemented, and thus codification is worthwhile.[117]

Certainly the shift in the standard of consideration, from "utmost deference" to "substantial weight," given to governmental assertions of the Privilege is the most substantive development.[118] The majority report on the bill notes that, "the government's assertions deserve weight and respect, but they do not deserve a reprieve from the rigorous, independent judicial scrutiny demanded by our adjudicatory system."[119]

Unquestionably, these are significant improvements towards ensuring that courts approach a government assertion of the Privilege with both greater skepticism and oversight. However, this legislation fails to protect inventors as well as it could, given that the "substantial weight" standard still greatly favors the Privilege. More problematically, the bill offers no recourse to inventors once the court concludes that material is privileged. Indeed, the majority report asserts that the bill provides a complete bar to

---

Agency, is reviewed in camera, and asserting "[h]owever helpful to the court the informed advocacy of the plaintiffs' counsel may be, we must be especially careful not to order any dissemination of information asserted to be privileged state secrets"); Jabara v. Kelley, 75 F.R.D. 475, 486-87 (E.D. Mich. 1977) ("In the case of claims of military or state secrets' privilege, however, the superiority of well-informed advocacy becomes less justifiable in view of the substantial risk of unauthorized disclosure of privileged information.").

115. S. Rep. No. 110-442, at 19.

116. *Id.* at 21.

117. *See id.* at 11 ("Many of these powers are already available to courts, but they often go unused.").

118. *Id.* at 51. Section 4054 lists the procedures for determining whether evidence is protected from disclosure by the state secrets privilege: "The court shall give substantial weight to an assertion by the United States relating to why public disclosure of an item of evidence would be reasonably likely to cause significant harm to the national defense or foreign relations of the United States." *Id.*

119. *Id.* at 12.

the use of privileged items in discovery or as evidence.[120] Why similar procedural mechanisms considered sufficient to resolve the Privilege issue could not be applied to the underlying litigation is unexplained by the majority report. By using some of those mechanisms, inventors' claims could proceed without risking national security in some situations. For instance, permitting the case to proceed *in camera*, with the documents kept under seal, would eliminate the concern that the information would be revealed as part of the public record.[121] If the government expresses concern that the information could still be too easily accessed, the court could add restrictions such as especially secure storage and encryption. Further, the entire proceeding could be held at a closed venue.[122] With regards to any concerns about appropriate personnel resource, private experts with the necessary clearances could be sought.[123]

The same security checks that could authorize opposing counsel for the Privilege hearing would presumably permit those attorneys to have access to try the case.[124] Alternatively, the court could order the govern-

---

120. *Id.* at 26.

121. Loral Corp. v. McDonnell Douglas Corp., 558 F.2d 1130, 1132-33 (2d Cir. 1977) (upholding both the trial judge's decision to refer the case to a magistrate for confidential proceedings, and requiring the defense department to provide the necessary security clearances for "the judge and magistrate assigned to the case, the lawyers and any supporting personnel whose access to the material is necessary" in an ordinary contract dispute concerning work performed on classified equipment for the Air Force); Halpern v. United States, 258 F.2d 36, 43 (2d Cir. 1958) ("[T]he privilege . . . is inapplicable when disclosure to court personnel in an in camera proceeding will not make the information public or endanger the national security.").

122. Clift v. United States, 597 F.2d 826, 829 n.2 (2d Cir. 1979) ("It would seem quite possible to have an [i]n camera production in this case in a secured area at Fort Meade, where the documents are."). Security-cleared court personnel could be provided by the executive department involved. *Id.* at 829 (suggesting that "some of the [g]overnment's objections to an [i]n camera trial in that case may have been exaggerated"). The court noted that "[i]t should not be difficult to obtain a court reporter and other essential court personnel with the necessary security clearance. If necessary, the stenographers who are now writing letters concerning this invention for the Department of the Navy can be utilized to record the testimony." *Id.* at 829 n.2.

123. *Id.* at 829 ("On the other hand while we sympathize with the judge's admission that she would be unable to understand the significance of the documents without the aid of an independent expert, efforts could be made to locate such an expert with appropriate clearances.").

124. *See* Section IV and notes 122-123. *But see* Halkin v. Helms, 598 F.2d 1, 3 (D.C. Cir. 1978) (rejecting the possibility that plaintiffs' counsel be present when the privileged information, interceptions of plaintiffs' foreign communications by the National Security Agency, is reviewed in camera). The court noted that "however helpful to the court the informed advocacy of the plaintiffs' counsel may be, we must be especially careful not to

ment to provide a plaintiff with suitable counsel. While few military attorneys have the requisite patent litigation background, the relative scarcity of these cases should make this option feasible. Plainly, this option would be less desirable; inventors would obviously prefer to choose their own counsel. Nonetheless, this alternative is likely to be more palatable to plaintiffs than to be stuck with unsupportable claims as a result of the suppression of crucial information.

One critical determinant as to whether these safeguards would be sufficient is the type of alleged invention at issue. In some cases, disclosure of the government's use of such devices to security-cleared litigants would appear unlikely to jeopardize national security. As Judge Newman noted, there may be areas of extreme sensitivity that cannot risk any judicial exposure (such as the Manhattan Project). A situation similar to the hypothetical disclosure of the Manhattan Project did, in fact, arise in *Hudson River Sloop Clearwater, Inc. v. Department of Navy*.[125] In *Hudson River Sloop*, more than ten environmental groups filed suit, asserting that the Navy had failed to engage in the proper environmental protection analysis prior to docking nuclear-armed vessels in New York ports.[126] The Navy claimed that the identity of nuclear-armed (as opposed to conventionally armed) ships was confidential, and a judicial decision in favor of the plaintiffs would breach national security by indicating to observers which ships carried nuclear weapons.[127] As the outcome of the suit could not remain secret, the Navy argued, the proceeding itself posed the threat of a violation of national security.[128] The court agreed, acknowledging that while in camera proceedings might provide some protection against public scrutiny, other parts of judicial proceedings are public and could result in a confidentiality breach.[129] Although the court could have ordered that its final determination remain sealed and placed a gag order on the plaintiff groups, given the number of people involved, as well as the yes-or-no na-

---

order any dissemination of information asserted to be privileged state secrets." *Id.* at 7. *See also* Jabara v. Kelley, 75 F.R.D. 475, 481 (E.D. Mich. 1977) ("In the case of claims of military or state secrets' privilege, however, the superiority of well-informed advocacy becomes less justifiable in view of the substantial risk of unauthorized disclosure of privileged information.").

125. Hudson River Sloop Clearwater, Inc. v. Dep't of the Navy, No. CV-86-3292, 1989 WL 50794 (E.D.N.Y. May 4, 1989).

126. *Id.*

127. *Id.* at *5.

128. *Id.*

129. *Id.* at *3.

ture of the confidentiality (the ships did or did not carry nuclear weapons), such a ruling was arguably appropriate.

By contrast, for the Crater coupler and most other inventions, merely elevating the protections afforded to the evidence would suffice.[130] The Crater coupler was intended to be used on underground fiber optic cables; the location of such cables themselves is not generally known, and they are usually miles under the ocean. Moreover, knowledge that Crater had prevailed in litigation, thus establishing that the United States uses a particular kind of coupler to connect cables in secret locations on the ocean floor, is almost certainly useless to the nation's enemies. Details will vary from case to case, but the Privilege does not justify the quashing of all such litigation.

It is true that there may be circumstances in which the danger is too great to justify providing even security-cleared inventors access to the privileged information. Consider, for instance, the cryptographic device at issue in *Clift*.[131] Concealment of cryptography is so critical to intelligence gathering that any possibility of its disclosure, even a small, carefully controlled one, is a significant threat to national security. In light of this, a court might reasonably conclude that no amount of procedural safeguards might be considered sufficient to permit the government to prove its use of an alternative design. But, in other cases, the sensitive information at issue is often already acknowledged to be in the possession of the inventor. For instance, in *Halpern*, the plaintiff asserted his claims under the Invention Secrecy Act; the purported privileged information was the patent application which the inventor himself submitted.[132] Responding to the government's invocation of the Privilege, the court ordered the case to proceed in camera. The court found the government's argument that any use of the application would constitute an unacceptable risk to national security unconvincing because the inventor was only seeking evidentiary proof of the submission of his own application.[133] Likewise, where the government had appropriated the patented invention as its own, the essence of the government's device is presumably already known to the patent holder. Even if some minor modifications were made to the device, redaction to withhold

130. Crater Corp. v. Lucent Techs., Inc., 423 F.3d 1260, 1271 (Fed. Cir. 2005).

131. Clift v. United States, 597 F.2d 826, 829 (2d Cir. 1979).

132. Halpern v. United States, 258 F.2d 36, 44 (2d Cir. 1958).

133. *Id.*; *see Clift*, 597 F.2d at 829 (noting that, as in *Halpern*, "Mr. Clift also knows what his invention is; indeed the world knows it as a result of the lifting of the secrecy order and subsequent issuance of the patent," but concluding that nonetheless Halpern might have had access to other classified information that he did not already possess).

those changes should avert any concern, and in such cases, the government's protestations ring hollow.

In sum, there may be circumstances in which even the enhanced procedural safeguards described in the bill, if applied beyond the Privilege hearing to actual litigation, would be insufficient to defend governmental interests. In that small fraction of national security situations, the complete suppression of sensitive information will be warranted. However, courts should be obliged to consider whether the particular circumstances would permit it to address any legitimate security concerns using the aforementioned procedural mechanisms rather than the draconian response of suppressing the material completely. Neither the *Crater* court nor the proposed bill offers any explanation for the failure to apply these types of procedural mechanisms to the underlying action. Along with the constitutional and long-term national security problems caused by the destruction of claims, the courts' fundamental obligation to hear valid claims should compel them to consider these alternatives instead of simply rubber-stamping the government's request for an extreme response to the invocation of the Privilege.

V. CONCLUSION

The result in *Lucent v. Crater* is every inventor's nightmare: a valuable idea, stolen, with no recourse. Unfortunately, rather than carefully parse out possible alternative approaches to recognizing the Privilege, the Federal Circuit permitted the Privilege to defeat the inventors' claims. In the context of claims regarding unauthorized military use of inventions, this reaction has the potential to engender several serious problems.

As demonstrated in *Crater*, under some circumstances, the suppression may disregard inventors' constitutional rights. Where the government has misappropriated (directly or indirectly) a trade secret, the owner has a constitutional right to compensation for the taking. Suppression of the evidence needed to prove the owner's claim removes this constitutionally mandated remedy. Adding salt to the constitutional wound, any direct takings claim would generally fall victim to the same problem.

Equally problematic is that destruction of inventors' claims will likely fall disproportionately on smaller businesses. This result is alarming, as it directly conflicts with the Department of Defense's contention that smaller businesses will need to produce a substantial share of the next generation of defense technology in order to maximize the nation's security, and its stated goal of encouraging such businesses to consider the Defense Department as a potential customer. Smaller businesses, already facing sub-

stantial obstacles in breaking into the defense procurement market, lack the political influence and long-term institutional relationships to ensure that the military will properly license their technology. As a result, these businesses must rely primarily on legal protection, rather than political or economic clout, to vindicate their intellectual property rights. If the application of the Privilege strips these smaller inventors of this protection, they may be disinclined to pursue inventions sought after by the military. In essence, the nation's long-term defense strategy is being taken hostage by a short-term litigation strategy.

Given these grave problems, as well as the judicial system's fundamental obligation to hear valid claims whenever possible, it is incumbent upon courts to consider more carefully whether the termination of most such litigation is truly required, or whether, in the presence of appropriate procedural mechanisms, some innovators could be permitted to pursue their claims. Indeed, recently proposed legislation instructs courts to use a variety of protective procedural mechanisms to protect sensitive information during the determination of the application of the Privilege. Yet the legislation inexplicably also authorizes, without apparent further consideration, complete suppression of any information eventually deemed privileged. This result appears unnecessary, as most of those same mechanisms could be applied equally well to an entire litigation proceeding.

It is true that there still may be the rare situation in which either the particularly sensitive nature of the invention, or the strategic inferences suggested by a judgment, indicates that no procedural safeguards would be sufficient to protect defense interests. In such a circumstance, there may be no recourse other than to terminate the litigation regardless of the immediate cost to the inventors and long-term costs to other government goals. But in contrary situations, which are likely to be more common, courts should search for an approach that will protect sensitive information without forcing inventors to have to pay the price (literally and figuratively) for the government's use of the Privilege.

LOCK DOWN ON THE THIRD SCREEN: HOW WIRELESS CARRIERS EVADE REGULATION OF THEIR VIDEO SERVICES

By Rob Frieden†

TABLE OF CONTENTS

---

† Pioneers Chair and Professor of Telecommunications and Law, Pennsylvania State University.

I. INTRODUCTION

Wireless handsets increasingly deliver more than tetherless telephone calls, text messaging, and ringtones. Next generation networks, and the sophisticated handsets that access them, offer a variety of new information, communications, and entertainment ("ICE") services. Wireless handsets have begun to work more like mobile computers by providing consumers with a "third screen" alternative for accessing multimedia content also available from television sets and personal computer monitors.[1]

The wireless third screen has the potential to offer users mobile access to everything the Internet offers. However, it has become clear that most wireless carriers have a financial interest in steering subscribers to a more limited "walled garden"[2] of content and software applications. Walled gardens provide carrier-delivered content to subscribers in a user-friendly, expedited basis. For example, a wireless carrier might offer access to a preferred search engine or source of content by offering a one step process. Subscribers seeking to access non-preferred content typically would have to undertake several additional steps that may add time, complexity, inconvenience, and possibly higher cost in the determination of whether to seek alternatives to walled garden options. Subscribers have access to walled garden content available from the carrier, an affiliate of the carrier, or a third party that has secured preferential access to the carrier's subscribers in exchange for sharing revenues. Unlike computer terminals where users have easy and unfettered opportunities to exit the carrier's walled garden, wireless subscribers may experience difficulty and incur additional costs when departing from wireless carriers' walled gardens. For example, Apple Computer, Inc. and AT&T, the exclusive United States vendor for the iPhone, offer an apparently quite generous walled garden containing over 30,000 applications that provide subscribers a variety of new options for accessing

1. *See, e.g.*, Nick Wingfield, *Time to Leave the Laptop Behind*, WALL ST. J., Oct. 27, 2008, at R1.

2. *In re* Implementation of Section 6002(B) of the Omnibus Budget Reconciliation Act of 1993, 23 F.C.C.R. 2241, 2315 (2008).

content and services.[3] But computer users would object to any software or content access limitation, including the obligation to use a single carrier-supplied interface to acquire that software, if an Internet Service Provider ("ISP") or a wired telephone company tried to establish such restrictions. Likewise, television owners would not tolerate any limitation on the devices they attach for accessing live or prerecorded content. The limits that wireless carriers impose on handset use starkly contrast with what carriers in other industries impose. ISPs have limited control over users' computer terminals, and video programmers have no power to restrict how television owners use their receivers to access programming. In contrast, wireless carriers can and do impose substantial operational limitations on handset use. Some of these restrictions represent necessary safeguards in light of the fact that wireless handsets use radio spectrum and require a technical interface to access specified channels using a predetermined format. However, most restrictions result from carriers' efforts to recoup financial subsidies given for handsets sold in conjunction with new or renewed service agreements. Additionally, carriers seek strategic opportunities to squeeze out more revenues per subscriber, to prevent migration to the services of another carrier, and to keep subscribers within the confines of a carrier's walled garden.

In the United States, the Communications Act of 1934, as amended, establishes separate definitions[4] for telecommunications,[5] information,[6] and cable television services.[7] The FCC appears unable to apply two or more regulatory classifications to the same operator when it provides more than one category of service.[8] Converging technologies and markets all but guar-

---

3. *See* Apple, Inc., App Store and Applications for iPhone, http://www.apple.com/iphone/appstore/ (last visited April 27, 2009).

4. *In re* Fed.-State Joint Bd. on Universal Serv., 13 F.C.C.R. 11501, 11522 (1998) ("[T]he language and legislative history of [the Communications Act of 1996] indicate that the drafters . . . regarded telecommunications services and information services as mutually exclusive categories."); *see also* Vonage Holdings Corp. v. Minn. Pub. Utils. Comm'n, 290 F. Supp. 2d 993, 1000 (D. Minn. 2003) (applying the FCC's dichotomy).

5. 47 U.S.C. § 153(43), (44), (46) (2006).

6. 47 U.S.C. § 153(20) (2006).

7. 47 U.S.C. § 522(6), (7), (20) (2006).

8. Technological and market convergence increasingly makes it difficult for the FCC to assign services into mutually exclusive categories, a task it considers compulsory. "[T]he language and legislative history of [the Communications Act of 1996] indicate that the drafters . . . regarded telecommunications services and information services as mutually exclusive categories." Federal-State Joint Board on Universal Service, Report to Congress, 13 F.C.C.R. 11501, 11522 (1998); *see also* Vonage Holdings Corp., 290 F. Supp. 2d at 994, 1000 (applying the FCC's dichotomy). "In keeping with the legislative history of the Communications Act, the Commission interprets that Act's definitions of 'telecommunications service' and 'information service' to be mutually exclusive." Communications Assis-

antee that operators will seek to expand their array of services to include both regulated telecommunications and other less regulated video and information services.[9] For example, a wireless carrier that provides conventional telephone services, treated as a regulated telecommunications service, also may offer Internet access, treated as a largely unregulated information service, and video services that trigger different types of regulations than telephone services. The Federal Communications Commission ("FCC") has evidenced an inability to apply two different regulatory regimes to the same venture when it offers services that trigger more than one regulatory classification, choosing instead to classify most ventures in the loosely regulated "information category." Accordingly, the FCC treats wireless carriers as qualifying for largely unregulated status, despite the fact that these carriers continue to provide regulated telecommunications services and offer video programming that could readily fit within the cable classification.

This Article will examine the regulatory status of wireless carrier-delivered video content with an eye toward determining the necessary scope and nature of government oversight. Past practices of the FCC in regulating the cable industry are examined to highlight the disparate treatment of wireless-delivered video content, despite the fact that both ventures distribute identical or similar content. The Article concludes that the FCC must balance the carriers' interests in finding new revenue centers to pay for next generation network upgrades with the subscribers' interests in having maximum freedom to use handsets they own. The Commission should help promote wireless carriers' exploitation of technological and marketplace innovations that make it possible to provide a combination of telecommunications, information, and video programming services. But regulators also

---

tance for Law Enforcement Act and Broadband Access And Services, ET Docket No. 04-295, RM-10865, First Report and Order and Further Notice of Proposed Rulemaking, 20 F.C.C.R. 14989, 14996 (2005) (citing Federal-State Joint Board on Universal Service, CC Docket No. 96-45, Report to Congress, 13 F.C.C.R. 11501, 11520, 11522-23 ¶¶ 39, 43 (1998)).

9. "Over the last two decades, the communications industry has undergone rapid technological advancements leading to the convergence of services. New technological capabilities allow companies to compete in markets which previously had no competition. While potentially beneficial to the consumer, convergence within the communications industry has created a regulatory nightmare." Ryan K. Mullady, *Regulatory Disparity: The Constitutional Implications of Communications Regulations That Prevent Competitive Neutrality*, 2 U. PITTSBURGH J. TECH. L. & POL'Y 4 (2007). For background on the impact of converging telecommunications and information processing technologies see, for example, International Telecommunication Union, ITU Internet Report 2006, digital.life; *available at* http://www.itu.int/osg/spu/publications/digitalife/index.html.

must guard against anticompetitive strategies designed to favor carrier ICE services by handicapping access to alternative sources. The Article concludes that because the FCC considers it necessary to promote video programming competition and access for wired cable television ventures, the Commission must undertake similar efforts to promote wireless access because wireless carriers have equivalent incentives and capabilities to blunt competition.

II. THE REGULATORY QUANDARY PRESENTED BY THIRD SCREEN CONVERGENCE

The versatility of modern wireless handsets presents a regulatory quandary for the FCC. Innovations in wireless handsets make it possible for devices to provide access to a variety of ICE services that include telephone calls, information services, e-commerce applications, position location functions, as well as access to the Internet and video programming. Wireless handsets have become an electronic Swiss Army knife capable of exploiting ICE convergence and easily toggling between first, second, and third generation wireless functions. In the first generation of wireless handsets, users made mobile telephone calls almost exclusively. With digitization in the second generation, subscribers could engage in text messaging, photography, music downloading, and other functions that rely on memory storage, keypads, and video screens. In the evolving third generation,[10] the wireless handset can switch between legacy functions, such as conventional voice telephony, and new features, such as broadband Internet access, that can convert the handset into a versatile platform for accessing most multimedia content.[11]

The FCC appears ill-equipped to apply different regulatory regimes to third generation wireless ventures capable of shifting functions. Prior to the onset of robust technological and market convergence, the FCC could enact different and mutually exclusive regulations based on the single set of services any one enterprise offered. The Commission established different regulatory requirements for broadcasters, cable television operators, telephone companies, and ISPs based on the specific characteristics of each type operator. The FCC has yet to address the impact of convergence that makes it

---

10. *In re* Serv. Rules for Advanced Wireless Servs. in the 1.7 GHz & 2.1 GHz Bands, 18 F.C.C.R. 25162, 25163 n.1 (2003) (describing the different generations of wireless technology).

11. Ed Rosenberg, Nat'l Regulatory Research Inst., Assessing Wireless and Broadband Substitution in Local Telephone Markets, (2007), http://nrri.org/pubs/telecommunications/07-06.pdf.

possible for wireless carriers to offer a combination of services, accessed by a single wireless device, that run the gamut of these regulatory classifications.

To make matters even more complicated in the wireless marketplace, vertically integrated wireless carriers blend content and conduit, making it all but impossible to establish "bright line" regulatory demarcations between the provision of telecommunications transmission services and the content these links transmit. In previous regulatory regimes, the FCC could separate the content provider from the content carrier, subjecting the former to little, if any, government oversight while subjecting the latter to extensive common carrier price, quality of service, and nondiscrimination regulation.

Using congressionally crafted service definitions,[12] the FCC applies the largely unregulated information services category to as many convergence services as plausible with an eye toward promoting marketplace-driven competition and innovation largely free of government oversight.[13] Traditionally, "information services" would be applied to providers of multimedia services while the heavily regulated "telecommunications" category would be applied to wireless carriers. The FCC has opted to ignore or subordinate the conduit function provided by wireless carriers providing multimedia services, which if separately identified would trigger the obligation for the Commission to enforce the more extensive regulations applied to telecommunications service providers. When wireless carriers offer a convergent blend of telecommunications access to telephony, video content, and information services, the FCC emphasizes the multimedia service delivered to subscribers' wireless handsets and deemphasizes the fact that the wireless carriers use a telecommunications network to deliver such content.[14] Reviewing courts have endorsed this regulatory and semantic sleight of hand.[15]

The FCC, while evidencing a preference to treat convergent services as "information services," has inconsistently applied legislatively crafted definitions with an eye toward exempting some convergent services while regu-

---

12. *See supra* notes 5-6.

13. *See, e.g.*, *In re* Appropriate Regulatory Treatment for Broadband Access to the Internet Over Wireless Networks, 22 F.C.C.R. 5901 (2007).

14. *In re* Inquiry Concerning High-Speed Access to the Internet over Cable & Other Facilities, 17 F.C.C.R. 4798 (2002), *aff'd sub nom.* Nat'l Cable & Telecomms. Ass'n v. Brand X Internet Servs., 545 U.S. 967 (2005).

15. Nat'l Cable & Telecomms. Ass'n v. Brand X Internet Servs., 545 U.S. 967 (2005) (upholding the FCC's determination that cable modem provided Internet access constitutes an information service); *see also* Rob Frieden, *What Do Pizza Delivery and Information Services Have in Common? Lessons From Recent Judicial and Regulatory Struggles with Convergence*, 32 RUTGERS COMPUTER & TECH. L.J. 247 (2006).

lating others. The FCC gladly deregulated all forms of wireline and wireless broadband access, but has not extended this regulatory forbearance to Voice over the Internet Protocol ("VoIP")[16] telephony and Internet Protocol Television ("IPTV").[17] These services seamlessly blend telecommunications transmission links with information services, but the FCC appears inclined to avoid applying the unregulated information services classification, because these services compete directly with pre-existing ("legacy") common carrier telephone and television service. Rather than treat VoIP carriers with the same sort of regulatory forbearance it applies to wireless telephone service, and increasingly to wireline service, the FCC has saddled VoIP service with regulatory burdens that make VoIP service more like conventional telephone service, at the expense of reducing VoIP's competitive cost advantage.[18] VoIP service providers, which offer subscribers telephone calling access to the conventional wireline public switched telephone network ("PSTN"), must contribute to universal service funding,[19] reconfigure their service to provide wiretapping capabilities to law enforcement authorities,[20] provide caller location identification and emergency 911 access[21] and offer service to disabled users.[22] Despite extensive rhetoric about refraining from imposing regulation on both emerging technologies and competitive services,[23] the FCC imposed these traditional regulatory burdens, choosing not to

---

16. Voice over the Internet Protocol ("VoIP") offers voice communications capabilities, much like ordinary telephone service, using the packet switched Internet, for all or part of the link between call originator and call recipient.

17. Internet Protocol Television ("IPTV") offers access to video programming via the Internet. Users can download files that contain such content for subsequent viewing. Alternatively they can receive an online "stream" of video packets corresponding to an existing file, or a simulcast of "live" programming.

18. *See* Rob Frieden, *Neither Fish Nor Fowl: New Strategies for Selective Regulation of Information Services*, 6 J. TELECOMM. & HIGH TECH L. 373 (2008).

19. *In re* Universal Serv. Contribution Methodology, 21 F.C.C.R. 7518, 7538 (2006) (extending section 254(d) permissive authority to require interconnected VoIP providers to contribute to the USF), *reh'g denied, vacated in part on other grounds*, Vonage Holding Corp. v. FCC, 489 F.3d 1232 (D.C. Cir. 2007).

20. *In re* Commc'ns Assistance for Law Enforcement Act & Broadband Access & Servs., 20 F.C.C.R. 14989 (2005), *petition for rev. denied*, 451 F.3d 226 (D.C. Cir. 2006).

21. *In re* IP-Enabled Servs., 20 F.C.C.R. 10245 (2005), *petition for rev. denied*, 473 F.3d 302 (D.C. Cir. 2006).

22. *In re* IP-Enabled Servs., 22 F.C.C.R. 11275 (2007), *rev'd*, 22 F.C.C.R. 18319 (2007) (granting in part and denying in part waivers of the FCC order).

23. Telecommunications Act of 1996, Pub. L. No. 104-104, § 706, 110 Stat. 56, 153 (codified as amended 47 U.S.C. § 157 note (2006)) (requiring the FCC to encourage the deployment of advanced telecommunications capability to all Americans by using regulating methods that remove barriers to infrastructure investment); *In re* Appropriate Regulatory Treatment for Broadband Access to the Internet Over Wireless Networks, 22 F.C.C.R.

allow the marketplace to determine whether the considerable service discounts available from VoIP service providers outweigh the greater risk in an emergency and greater inconvenience for some users.

The FCC has largely accepted the view that wireless carriers need to qualify for a deregulated "safe harbor"[24] so that sufficient incentives exist for these carriers to invest in next generation infrastructure and spectrum auctions.[25] This preoccupation with incentive creation ignores the likely probability that incumbent wireless carriers would invest in any new spectrum to foreclose market entry by new competitors. The attractiveness of investing in wireless technology is illustrated by the growing reliance of incumbent wireline carriers on wireless services to generate revenues in light of declines in previous core market segments, e.g., local and long distance telephone service.

III. WIRELESS CARRIERS' INCENTIVE AND ABILITY TO LOCK DOWN ACCESS

The FCC's willingness to allow wireless carriers to deprive subscribers of the access freedoms mandated for wire-based carriers caters to the wireless carriers' large financial incentive to impose limits on handsets, and keen interest in avoiding rules that could impair their ability to maintain such limits. Wireless carriers have invested billions of dollars in the spectrum[26] and infrastructure needed to provide third generation network services. Additionally, these carriers typically combine wireless service with offers of a subsidized handset in exchange for a two year service commitment by subscribers. In light of such risk taking, these carriers predictably seek to evade, or at least limit, any government regulatory oversight that might constrain their ability to recoup their handset subsidies and network facilities

---

5901, 5911 (2007) (classifying wireless broadband Internet access as a lightly regulated information service to "help spur growth and deployment of these services"); *see also* 47 U.S.C. §§ 157, 230(b)(2) (2006) (stating that it is the policy of the United States "to preserve the vibrant and competitive free market that presently exists for the Internet").

24. BLACK'S LAW DICTIONARY 1363 (8th ed. 2004) (defining a safe harbor to be "[a]n area or means of protection [or a] provision (as in a statute or regulation) that affords protection from liability or penalty.").

25. *In re* Appropriate Framework for Broadband Access to the Internet Over Wireline Facilities, 20 F.C.C.R. 14853, 14878 (2005) ("[W]e seek to adopt a comprehensive policy that ensures . . . that broadband Internet access services are available to all Americans and that undue regulation does not constrain incentives to invest in and deploy the infrastructure needed to deliver broadband Internet access services").

26. *See* FCC, Auctions Summary, http://wireless.fcc.gov/auctions/default.htm?job=auctions_all (last modified Nov. 10, 2008).

investment, particularly by diversifying services and increasing the monthly average return per user ("ARPU").

However, the requirement that common carriers operate in a nondiscriminatory manner constitutes a fundamental component of conventional common carrier telecommunications service regulation and should apply to wireless carriers when they provide voice telephone services.[27] Title II of the Communications Act of 1934, as amended, requires telecommunications service providers, *inter alia*, to charge just and reasonable rates, to operate in a nondiscriminatory manner, and to submit to scrutiny by the FCC to resolve consumer complaints. In their capacity as information and video programming providers, wireless carriers do not trigger conventional telephone service common carrier regulation. The FCC can forbear from applying any of the common carrier regulations to wireless carriers only upon determining that consumers will remain protected against unreasonable and discriminatory practices and that the public interest supports forbearance.[28] Wireless carriers want the FCC and the public to conclude that even though wireless carriers still must comply with most Title II common carrier telecommunications service regulations,[29] the FCC nevertheless should refrain from enforcing the rules in light of the fact that the carriers also offer multimedia services which qualify for the largely unregulated "information services" category.

Wireless carriers should not qualify for exemption from common carrier regulation[30] simply because a portion of what they offer does qualify for comparatively less regulatory oversight than their telecommunications service. Nevertheless, a prominent *Wall Street Journal* industry analyst has concluded that the wireless carriers have succeeded in creating the inference that they cannot be regulated:

> A shortsighted and often just plain stupid federal government has allowed itself to be bullied and fooled by a handful of big wireless

---

27. The FCC uses the term Commercial Mobile Radio Service ("CMRS") to identify the basic telephone services provided by wireless carriers. *See* Omnibus Budget Reconciliation Act of 1993, Pub. L. No. 103-66, § 6002, 107 Stat. 312, 393 (codified as amended at 47 U.S.C. § 332 (2006)) (creating the CMRS carrier category).

28. 47 U.S.C. § 332(c)(1)(A)(i)-(iii) (2006); *see also* 47 U.S.C. § 160(a) (2006) (establishing similar forbearance criteria for other telecommunications service providers).

29. 47 U.S.C. §§ 201-276 (2006).

30. Despite having a statutory duty to regulate cellular telephone carriers as common carriers, the FCC rarely acknowledges and acts on this requirement. *See In re* Reexamination of Roaming Obligations of Commercial Mobile Radio Serv. Providers, 22 F.C.C.R. 15817, 15818 (2007) (specifying that cellular operators must provide their subscribers automatic access to other carriers to for making and receiving telephone calls when traveling outside subscribers' home service regions).

> phone operators for decades now. And the result has been a mobile phone system that is the direct opposite of the PC model. It severely limits consumer choice, stifles innovation, crushes entrepreneurship, and has made the U.S. the laughingstock of the mobile-technology world, just as the cellphone is morphing into a powerful hand-held computer . . . . That's why I refer to the big cellphone carriers as the "Soviet ministries." Like the old bureaucracies of communism, they sit athwart the market, breaking the link between the producers of goods and services and the people who use them.[31]

In this light, the FCC all the more needs to remedy carriers' tactics that are detrimental to market competition. Instead, the FCC strongly prefers to shoehorn any and all converged services into the lightly regulated information services "safe harbor," including wireless broadband Internet access service.[32] With rare exception, the FCC appears reluctant to hold wireless operators to the still applicable Title II requirements, despite having not untaken the examination necessary to forbear officially from regulating.[33] The FCC has evidenced an inability to subject wireless carriers to multiple regulatory regimes based on the fact that these carriers trigger different regulatory scrutiny when providing different services. Without undertaking the appropriate administrative process to forbear from regulating wireless carriers under Title II, the FCC simply refrains from providing the necessary scrutiny. Regardless of the regulatory posture undertaken by the FCC, or what regulatory classification the FCC applies to particular wireless services, fundamental public interest and consumer protection concerns require the FCC to assess whether wireless carriers have the incentive and ability to discriminate unlawfully and to regulate appropriately. Wireless carriers' efforts[34] to frame the debate in terms of whether their broadband and video

---

31. Posting of Walt Mossberg to All Things Digital (Mossblog), Free My Phone, http://mossblog.allthingsd.com/20071021/free-my-phone/ (Oct. 21, 2007, 21:31 PT).

32. *See In re* Appropriate Regulatory Treatment for Broadband Access to the Internet Over Wireless Networks, 22 F.C.C.R. 5901 (2007).

33. 47 U.S.C. § 160(a)(1)-(3) (2006) (authorizing the FCC to forbear from applying specific aspects of Title II regulation if enforcement only upon determining that consumers will remain protected against unreasonable and discriminatory service and that the public interest supports forbearance); *see also*, MCI Telecomms. Corp. v. Am. Tel. & Tel. Co., 512 U.S. 218 (1994) (absent explicit legislative authorization the FCC could not exempt designated "nondominant" carriers from the Communications Act's Title II requirement that all carriers file a public contract, known as a tariff, specifying the terms and conditions of service).

34. *In re* Appropriate Regulatory Treatment for Broadband Access to the Internet Over Wireless Networks, 22 F.C.C.R. 5901, 5905-06 (2007) ("Cingular Wireless & BellSouth and Cisco contended that wireless broadband Internet access service should be classi-

programming services qualify for deregulation as information services largely shifts attention from the real problem, which is whether government oversight remains legislatively mandated and necessary, or whether a fully competitive and self-regulating marketplace exists. In the following Parts, the Article will draw analogies from the FCC's regulation of similar industries to show that regulation of the wireless carriers remains consistent with FCC policy and practice. When cable television ventures attempt to favor corporate affiliates, or stifle competition in the delivery of, and access to, video programming, the FCC has aggressively intervened.

IV. CURBING CABLE TELEVISION MARKET POWER IN VIDEO PROGRAMMING ACCESS

The FCC's aggressive and expanding regulation of cable companies stands in sharp contrast with its reluctance to enforce valid regulations of wireless carriers, despite the similarity of content provided to consumers. For over several decades, the FCC has expressed concerns that cable television ventures can and will harm consumers by establishing vertically integrated firms[35] that create and distribute "must-see" content, while limiting access to multi-channel video programming distributor ("MVPD") competitors. The major cable television companies both create and distribute video programming. The FCC attempts to restrict the power of cable companies which have every financial incentive to constrain consumers' choice of video programming available from competitors.

In 2007, the FCC specifically found that cable companies tied programming and distribution in ways that reduced the programming available to MVPD competitors:

> [W]e conclude that there are no good substitutes for some satellite-delivered vertically integrated programming and that such

---

fied as an information service, asserting that the deregulatory features of CMRS allowed under section 332 were not sufficient.").

35. Vertical integration refers to the combination of separate market activities by a single enterprise. *In re* Annual Assessment of the Status of Competition in the Market for the Delivery of Video Programming, 21 F.C.C.R. 2503, 2575 (2006) ("Vertical relationships may have beneficial effects, or they may deter competitive entry in the video marketplace and/or limit the diversity of programming."); *id*. at 2575 n.565 ("Beneficial effects can include efficiencies in the production, distribution, and marketing of video programming, and providing incentives to expand channel capacity and create new programming by lowering the risks associated with program production ventures"); *id*. at 2575 n.566 ("Possible detrimental effects can include unfair methods of competition, discriminatory conduct, and exclusive contracts that are the result of coercive activity.").

> programming therefore remains necessary for viable competition in the video distribution market. Based on this finding, we conclude that vertically integrated programmers continue to have the ability to favor their affiliated cable operators over competitive MVPDs such that competition and diversity in the distribution of video programming would not be preserved and protected absent the rule. Although we find some trends in the markets for both video programming and video distribution since 2002 that might decrease the incentive of vertically integrated programmers to withhold programming from competitive MVPDs, we also find some trends that increase their incentive to withhold programming, such as the increase in horizontal consolidation of the cable industry, the increase in cable clustering, and the recent emergence of new competitors. We also find specific factual evidence that, where the exclusive contract prohibition does not apply, such as in the case of terrestrially delivered programming, vertically integrated programmers have withheld and continue to withhold programming from competitive MVPDs.[36]

Because cable television companies generate most of the desired video content and control the major medium for distributing that content, the FCC expressed concern that the cable companies can thwart competition, favor affiliated content providers, stifle the development of new content sources and extract rates above competitive levels from subscribers.[37]

A. Ban on Exclusive Contracts between Programmers and Cable Operators

Troubled by the power of cable television ventures to stifle consumer access to content, the FCC decided to extend a ban on exclusive contracts between vertically integrated programmers and cable operators for five additional years, until October 5, 2012.[38] The FCC determined that vertically integrated programmers still have the ability[39] and the incentive[40] to favor

---

36. *In re* Implementation of the Cable Television Consumer Protection & Competition Act of 1992, 22 F.C.C.R. 17791, 17810 (2007).

37. *Id*. at 17816 (expressing concern over vertically integrated cable operators).

38. *Id*. at 17792 (finding that the exclusive contract prohibition was necessary "to preserve and protect competition and diversity in the distribution of video programming").

39. *Id*. at 17814-15 (expressing the commission's concern that the four largest cable MSOs have an interest in the top networks as ranked by subscribership).

40. *Id*. at 17821 (describing how an exclusive arrangement between a cable-affiliated programmer and its affiliated cable operator will increase the number of subscribers when customers switch to the affiliated able distribution service in order to receive the exclusive programming).

corporate affiliates over other competitive providers.[41] The FCC reaffirmed its conclusion that vertically integrated ventures still control the "must see" content, for which no viable substitute exits.[42] The FCC retained the prohibition against exclusive content distribution contracts from ventures that vertically integrate content production and distribution to consumers.

The FCC showed strong resistance to vertically integrated carriers' control of content provided to users. The FCC rejected programmer suggestions to narrow its ban on exclusive contracts so that restrictions would vary with the popularity of the programming network and the competitive climate in specific regions served by a cable operator.[43] Additionally the FCC refused to limit the restriction to conventional cable television operators, which would exclude other MVPDs, or to cable operators that have been in the MVPD market for more than five years, have extensive resources, and have entered into exclusive contracts for programming.[44]

In contrast to the FCC's strong preference to regulate the cable industry, the FCC declined to expand the exclusive contract prohibition to apply to non-cable-affiliated programming, e.g., content created by vertically integrated direct broadcast satellite ("DBS") operators and new MVPDs such as AT&T and Verizon. The FCC also concluded that section 628(c)(2)(D) did not apply to terrestrially delivered programming because specific statutory language has limited the exclusive contract prohibition to content delivered via satellite. However, in light of finding that a vertically integrated cable television operator had withheld terrestrially delivered regional sports network content in San Diego and Philadelphia, the FCC nevertheless considered whether to extend the program access rules to all terrestrially delivered cable-affiliated programming.[45] While the FCC has shown great concern about consumer access to content controlled by cable operators, it has no such concern that vertically integrated telephone companies would secure exclusive content distribution agreements.

B. Prohibition on Content Tier-Buy Through Requirements

The FCC limits the ability of cable operators to foist unwanted and costly content on subscribers as prerequisites for accessing desired content.

---

41. *Id.* at 17819 (finding that access to vertically integrated programming is essential for new entrants in the video marketplace to compete effectively).

42. *Id.* at 17810 ("[W]e conclude that there are no good substitutes for some satellite-delivered vertically integrated programming and that such programming therefore remains necessary for viable competition in the video distribution market.").

43. *Id.* at 17841 (declining proposals to narrow the ban on exclusive contracts).

44. *Id.* at 17842.

45. *Id.* at 17860.

Section 3 of the 1992 Cable Act[46] prohibits cable television operators, operating in a market without effective competition, from requiring subscribers to "buy through"[47] intermediate tiers of programming in order to access desired content in a higher service tier. Thus, consumers do not have to subscribe to the so-called enhanced basic service tiers, which bundle a variety of cable television programming, before they can view content offered on a per view or per channel basis, such as individual premium channels like Home Box Office.

Former FCC Chairman Kevin Martin promoted the option of content selection on an à la carte, network-by-network basis in lieu of service tiers that contain many channels of content, some of which individual consumers consider offensive.[48] In contrast to these limits imposed by the FCC on cable television providers, the FCC imposes no regulations on how wireless carriers package and price access to video content.

C. Forced Alternatives to Mandatory Set Top Box Leases

The FCC has designed rules to enable cable television subscribers to access content via "cable ready" television sets[49] without having to lease a device, known as a set top converter, that is necessary to descramble signals. The FCC prohibits cable television companies from offering set top converters that combine security functions, e.g., descrambling, with other features, such as channel selection and navigation, electronic program guides, and pay per view, on-demand access to content. This prohibition prevents cable companies from requiring all subscribers to lease set top boxes.[50] With the prohibition, cable television companies can perform secu-

---

46. Cable Television Consumer Protection and Competition Act of 1992, Pub. L. No. 102-385, § 3, 106 Stat. 1460, 1464 (codified as amended at 47 U.S.C. § 543(80)(A) (2006)).

47. *In re* Sections of the Cable Television Consumer Prot. & Competition Act of 1992, 9 F.C.C.R. 4316, 4327 (1994) ("The tier buy-through prohibition of the 1992 Cable Act prohibits cable operators from requiring subscribers to purchase a particular service tier, other than the basic service tier, in order to obtain access to video programming offered on a per-channel or per-program basis."); *see also* FCC Consumer Options for Selecting Cable Channels and the Tier Buy-Through Prohibition, Feb. 25, 2003, *available at* http://fjallfoss.fcc.gov/edocs_public/attachmatch/DOC-231469A1.pdf (FCC Fact Sheet).

48. Letter from Kevin Martin, Chairman, FCC, to Gary Flowers et al. (Aug. 22, 2007), *available at* http://www.fcc.gov/commissioners/previous/martin/alacarte-ltr-082207.pdf (letter to representatives of several minority public advocacy organizations).

49. *In re* Implementation of Section 304 of the Telecomms. Act of 1996, 20 F.C.C.R. 6794 (2005).

50. *Id*. at 6807-08.

rity and digital rights management via a computer chip known as a CableCard that subscribers insert into most recent vintage television sets.[51]

Several cable operators unsuccessfully challenged the FCC's prohibition on set top boxes providing both security and non-security functions. On two occasions,[52] the D.C. Circuit affirmed the FCC on several grounds. The court refused to consider petitioners' statutory claim that a difference exists between set top converter boxes and other equipment within the context of section 629(a) of the Communications Act, which states that the FCC "shall not prohibit any [MVPD] from also offering converter boxes, interactive communications equipment, and other equipment used by consumers to access multi-channel video programming."[53] The court characterized the petitioners' claim as illogical: "[I]f integrated set-top boxes are not 'converter boxes,' as we held in *General Instrument*, then they must be 'other equipment,' a possibility we did not address there. And if integrated boxes are 'other equipment,' then section 629(a)'s second sentence prevents the FCC from barring cable operators from offering them."[54]

The court refused to consider this statutory claim on two procedural grounds: (1) section 629(a) established a sixty-day time period for any petitions for review of applicable Commission orders; and (2) the petitioners never presented this issue for consideration by the FCC and therefore section 405 of the Communications Act precludes raising the issue on appeal. The court upheld the FCC's integration requirement against the petitioners' argument for abandoning the requirement. While CableCard compatible television sets had become commonplace, few consumers used CableCards. The court held that "there was nothing unreasonable about the FCC's conclusion that 'the competitive reasons that led the Commission to impose the integration ban have not been eliminated by the developments in the market.' "[55]

The court also rejected the claim that the FCC arbitrarily exempted DBS operators from the integration ban. The court upheld the exemption based on the FCC's interpretation of the statutory criteria identifying instances

---

51. Charter Commc'ns, Inc. v. FCC, 460 F.3d 31 (D.C. Cir. 2006) ("[A] CableCARD . . . plugs into a slot in a host navigation device, permitting the device to perform both the security and non-security functions.").

52. *See id.*; Gen. Instrument Corp. v. FCC, 213 F.3d 724 (D.C. Cir. 2000) (affirming the FCC's statutory authority to require separation of security and other set top converter functions); *In re* Implementation of Section 304 of the Telecomms. Act of 1996, 13 F.C.C.R. 14775 (1998).

53. 47 U.S.C. § 549(a) (2006).

54. *Charter Commc'ns*, 460 F.3d at 37-38.

55. *Id.* at 41 (citing *In re* Implementation of Section 304 of the Telecomms. Act of 1996, 20 F.C.C.R. 6794 (2005)).

where an MVPD allows alternatives to its set top box: a DBS is exempt when an MVPD "supports the active use by subscribers of navigation devices that: (i) operate throughout the continental United States, and (ii) are available from retail outlets . . . throughout the United States that are not affiliated with the [MVPD]."[56] The court noted that DBS operators have met the requirements while "the vast majority of cable subscribers remain dependent upon non-portable converter boxes available only from their cable companies."[57] In addition, the court rejected the cable operators' claims that increased facility-based competition, e.g., video program delivery from telephone companies, has encouraged cable companies to offer consumers every possible equipment alternative.

> [W]hatever the theoretical incentives, the FCC found that the real-world result that section 629(a) commanded it to assure-the commercial availability of navigation devices from vendors unaffiliated with MVPDs-has not arrived.[58]

Additional objections about unforeseen costs that cable companies would incur failed to persuade the court:

> The Commission also took steps to minimize industry costs, both by extending the implementation deadline from 2006 to 2007, and by promising to reconsider eliminating the ban altogether should the cable and consumer electronics industries achieve a downloadable security solution capable of providing common reliance without requiring the physical separation of security and non-security functions.[59]

Cable operators have largely thwarted the Congressional mandate to give consumers alternatives to the operator-leased devices. However, while a competitive market for such devices has not evolved and few consumers even know about the CableCard option, recent innovations in digital video recorders may incorporate many of the features provided by the cable operators. Recent FCC decisions suggest that cable operators will no longer succeed in stalling compliance with section 629 of the Communications Act.[60]

---

56. 47 C.F.R. § 76.1204(a)(2) (2005).
57. *Charter Commc'ns*, 460 F.3d at 43.
58. *Id*. at 44.
59. *Id*. at 42
60. *See In re* Implementation of Section 304 of the Telecomms. Act of 1996, 20 F.C.C.R. 6794 (2005); *see also In re* Comcast Corp., 22 F.C.C.R. 228 (2007).

V. REASONS WHY WIRELESS CARRIERS SHOULD BEAR CONTENT NONDISCRIMINATION REQUIREMENTS LIKE THOSE BORNE BY CABLE OPERATORS

Given wireless carriers' practice of marketing handsets bundled with service, the concentrated nature of the industry, and the rewards accruing from walled gardens, the FCC should subject wireless carriers to the same sort of nondiscrimination rules applied to cable operators. Because wireless handsets likely will become a competitive alternative to televisions and computer terminals, regulatory parity requires that vertically integrated wireless ventures bear similar burdens as cable operators if both type of ventures have similar incentives and abilities to exploit market power.

A. Public Benefits and Harms from Bundling Wireless Service with Handsets

United States wireless carriers can restrict the versatility and utility of handsets primarily because most subscribers acquire subsidized handsets in exchange for a two-year service commitment[61] and carrier-imposed contractual limitations on handset use. Wireless subscribers have grown accustomed to acquiring "free" or inexpensive handsets. Perhaps this benefit offers ample compensation even though wireless service rates factor in the cost of the subsidy and subscribers incur substantial financial penalties for changing carriers before completion of the two-year service commitment.[62]

However, wireless carriers can reduce the versatility and the utility of the handset by locking users' access to content and locking out content providers' access to users. Here is a list of existing lock-out strategies:

> Locking handsets so that they cannot access competitor networks (by frequency, transmission format, firmware or software). In the U.S. carriers even lock handsets designed to allow multiple carrier access by changing an easily inserted Subscriber Identity Module ("SIM");
>
> Using firmware "upgrades" to "brick," i.e., render inoperative, the handset or alternatively disable third-party hardware and software;
>
> Disabling handset functions, e.g., Bluetooth, Wi-Fi access, Internet browsers, GPS services, and email clients;

---

61. *In re* Implementation of Section 6002(B) of the Omnibus Budget Reconciliation Act of 1993, No. 08-27, 2009 WL 151633, ¶ 113 (Fed. Commc'ns Comm'n 2009).

62. The fixed-term service contracts and ETFs are traditional tactics of the industry to provide a discount upfront and increase sales. *Id.*

> Specifying formats for accessing memory, e.g., music, ringtones, and photos;
>
> Creating "walled garden" access to favored video content of affiliates and partners; and
>
> Using proprietary, non-standard interfaces difficult for third parties to develop compatible applications and content.[63]

The FCC needs to restrain wireless carriers from imposing limitations on handsets that have nothing to do with legitimate network management and everything to do with favoring affiliated content providers over third party service providers. These safeguards should stipulate that subscribers have a right to use any technically compatible handset to access any available source of content, software, or computer application whether or not affiliated with the wireless carrier providing the link.[64]

Long ago, the FCC rejected any attempt by wireline carriers to limit, block, or disable access by handsets bought from unaffiliated suppliers.[65] Consumers take for granted the right to buy and operate their own telephones for accessing wireline networks, and the same principle should ap-

---

63. AT&T, Wireless Data Service Terms and Conditions, http://www.wireless.att.com/learn/messaging-internet/media-legal-notices.jsp (last visited Apr. 22, 2009) (listing prohibited uses under its plan and the consequences of violations); *see also* AT&T, Acceptable Use Policy, http://www.corp.att.com/aup/ (last visited Apr. 22, 2009). The policy spells out a range of activities:

> AT&T prohibits use of the IP Services in any way that is unlawful, harmful to or interferes with use of AT&T's network or systems, or the network of any other provider, interferes with the use or enjoyment of services received by others, infringes intellectual property rights, results in the publication of threatening or offensive material, or constitutes Spam/E-mail/Usenet abuse, a security risk or a violation of privacy.

Id.

But see Letter from Ben Scott, Policy Dir. & Chris Riley, Policy Counsel, Free Press, to Michael J. Copps, Acting Chairman, FCC (Apr. 3, 2009), *available at* http://fjallfoss.fcc.gov/prod/ecfs/retrieve.cgi?native_or_pdf=pdf&id_document=6520205185 (opposing AT&T decision to block wireless subscribers from activating Skype VoIP software when using the AT&T network, but allowing such software when subscribers have access to Wi-Fi networks (ex parte communication regarding WC Docket No. 07-52)).

64. Rob Frieden, *Hold the Phone: Assessing the Rights of Wireless Handset Owners and Carriers*, 69 U. PITT. L. REV. 675, 720-25 (2008).

65. *See* Hush-a-Phone v. United States, 238 F.2d 266, 269 (D.C. Cir. 1956) (ordering the FCC to eliminate telephone company tariff restrictions on customers' right to attach non-electronic acoustic devices to telephones). In 1968 the FCC extended the right to include attachment of electronic devices. *In re* Use of the Carterfone Device in Message Toll Tel. Serv., 13 F.C.C.2d 420 (1968), *recon. denied*, 14 F.C.C.2d 571 (1968).

ply for access to wireless networks.[66] In the *Carterfone* policy of 1968, the FCC specifically mandated all wireline telephone companies to allow subscribers to attach any technically compatible device.[67]

Already many purchasers of Apple iPhones and other cellphones have resorted to self-help[68] tactics to eliminate manufacturer or carrier-imposed limitations on the handset's versatility, features and on access to third party applications and content. Rather than allow carriers to punish subscribers for engaging in such tactics, the FCC should establish a process to certify handset technology to allow any handset, operating in the proper format and frequency, to access any carrier's network and any software application. At the very least, the FCC should forbid wireless operators from imposing handset restrictions unrelated to legitimate network management.

B. The Wireless Carrier Marketplace is as Concentrated and Anticompetitive as Cable Television

The FCC recognizes that vertical integration in video content creation and distribution requires regulatory intervention. Commercial Mobile Radio

---

66. *See* Rob Frieden, *Wireless Carterfone—A Long Overdue Policy Promoting Consumer Choice and Competition* (New Am. Found., Wireless Future Program, Working Paper No. 20), *available at* http://www.newamerica.net/files/Wireless_Carterfone_Frieden.pdf.

67. The *Carterfone* policy established by the FCC in 1968 requires all wireline telephone companies to allow subscribers to attach any technically compatible device. Consumers take for granted the right to attach any device to a network that is "privately beneficial without being publicly harmful." *See Use of the Carterfone Device*, 13 F.C.C.2d at 423; *see also* Pub. Util. Comm'n of Tex. v. FCC, 886 F.2d 1325, 1329 (D.C. Cir. 1989) (noting long established FCC policy that carriers and non-carriers alike have a federal right to interconnect to the public telephone network in ways that are privately beneficial if they are not publicly detrimental); *In re* Am. Tel. & Tel. Co.'s Proposed Tariff Revisions, 53 F.C.C.2d 473, 477 (1975), *aff'd sub nom.* Mebane Home Tel. Co. v. FCC, 535 F.2d 1324, 1329 (D.C. Cir. 1976); *In re* Telerent Leasing Corp., 45 F.C.C.2d 204, 205 (1974), *aff'd sub nom.* N.C. Util. Comm'n v. FCC, 537 F.2d 787 (4th Cir. 1976).

Previous FCC opposition to this principle failed to pass muster with a reviewing court that interpreted the Communications Act as mandating the right of consumers to attach equipment to the network in ways that were privately beneficial but not publicly harmful. *Hush-A-Phone*, 238 F.2d at 266, 269 ("The intervenors' tariffs [prohibiting the use of plastic device to enhance privacy and low volume conversations], under the Commission's decision, are in unwarranted interference with the telephone subscriber's right reasonably to use his telephone in ways which are privately beneficial without being publicly detrimental.").

68. One such tactic is "jailbreaking" the phone, which allows a user to install on his device third-party applications unapproved by the provider. *See* Sarah Perez, *Why You Have To Jailbreak the iPhone*, N.Y. TIMES, Jan. 12, 2009, http://www.nytimes.com/external/readwriteweb/2009/01/12/12readwriteweb-why_you_have_to_jailbreak_the_iphone.html.

Service ("CMRS") operators, FCC's term for cellular radio telephone carriers, benefit from a similarly concentrated and vertically integrated industry. The top two CMRS carriers, AT&T and Verizon, control at least 53.4 percent of the wireless market,[69] and these companies also have substantial market share in broadband wireline access, e.g., Digital Subscriber Link[70] and wireline telephone service. In addition to the market power accruing from a commanding share of the wireless industry, AT&T and Verizon vertically integrate by securing exclusive content distribution rights for carriage via their wireless networks. They horizontally integrate by bundling triple-play[71] and quadruple-play service packages[72] combining wireless service with wireline telephony, Internet access, and wireline video program access. With such a broad wingspan of packaged services, these ventures can become commingle costs with an eye toward driving out competitors with a "one-stop shop" that can result in further industry consolidation through less competition and greater market shares held by a few firms.

C. Incentives and Rewards from Wireless Walled Gardens

Walled gardens can yield higher revenues for wireless carriers by greatly improving the odds that subscribers either will pay for carrier-supplied content, or will view advertising supported content that appears within the carrier's walled garden. When subscribers have to undertake several navigational steps and possibly incur higher costs to access content outside the carrier's walled garden, many will make do with what their carrier provides.

For subscribers using subsidized handsets, a wireless carrier can block access to competing carriers' networks and unaffiliated content sources. The carrier typically reserves the right to impose such access restrictions via the

---

69. Leslie Cauley, *iWeapon: AT&T Plans to Use its Exclusive iPhone Rights to Gain the Upper Hand in the Battle for Wireless Supremacy*, USA TODAY, May 22, 2007, at 1B. The top four carriers control 88.1 percent of the wireless telecommunications market and that figure will near ninety percent upon completion of Verizon's acquisition of Alltel.

70. Digital Subscriber Links provide Internet access via the copper wires initially used solely to provide narrowband telephone service. Telephone companies retrofit the wires to provide medium speed broadband services by expanding the available bandwidth by about 1500 kiloHertz. The FCC defines Digital Subscriber Line as a technology that brings high-speed and high-bandwidth information over existing telephone lines. FCC, Broadband Access for Consumers, http://www.fcc.gov/cgb/consumerfacts/dsl2.html (last modified Feb. 12, 2002).

71. *In re* Exclusive Serv. Contracts for Provision of Video Servs. in Multiple Dwelling Units & Other Real Estate Devs., 22 F.C.C.R. 5935, 5398 (2007) (notice of proposed rulemaking).

72. The quadruple play refers to the combination of "video, broadband Internet access, VoIP and wireless service." *In re* AT&T Inc. & Bellsouth Corp., 22 F.C.C.R. 5662, 5735 (2007).

non-negotiable subscription agreement, which few customers read and understand. The imposition of restrictions and financial penalties for early termination of service shows that the offer of a "free" wireless handset has many strings attached and in application constitutes an installment-sales contract containing substantial restrictions on subscriber use. Yet, even for subscribers who fully own their handset, whether by completing the initial two years of "rent to own" period, or acquiring an unsubsidized unit, the FCC does not outlaw restrictions that foreclose wireless subscribers freedoms to access preferred content and software applications.

VI. THE RELEVANCE OF THE INTERNET FOUR- FREEDOMS POLICY IN THE *COMCAST* ORDER

The lack of any handset attachment freedoms for wireless access contrasts with an explicit FCC endorsement of the *Carterfone* policy in 2005 as an integral part of what Internet access rights consumers should have. The FCC should extend regulation to wireless carriers in accordance with its Internet Four-Freedoms Policy. The FCC articulated four essential "Internet Freedoms" emphasizing nondiscrimination in the manner by which ISPs provide Internet access to end users and the delivery of content.[73] The FCC expects all ISPs, regardless of corporate affiliation and technology, to provide subscribers with the freedom to: (1) access the lawful content of their choice; (2) run applications and services of their choice, subject to the needs of law enforcement; (3) connect their choice of legal devices that do not harm the network; and (4) benefit from competition among network providers, application and service providers, and content providers.[74]

A. Applying the Four Freedoms to Network Interference

Even though the FCC did not formalize 2005 articulation of Internet Freedoms as rules, it considers it an enforceable policy and has in fact enforced this policy against Comcast. The Commission decided that Comcast violated this policy when the company used network management techniques to choke peer-to-peer ("P2P")[75] transfers even when the company experienced no congestion and could have handled the traffic without imposing delays. By a 3-2 vote, the FCC concluded that Comcast violated the

---

73. *In re* Appropriate Framework for Broadband Access to the Internet over Wireline Facilities, 20 F.C.C.R. 14986 (2005) [hereinafter 2005 Internet Policy Statement].

74. *Id.* at 14988.

75. Peer-to-peer file transfers involves the transfer of content between two or more participants. With the use of software, such as BitTorrent, files can arrive in a timely manner by subdividing content into smaller units that are routed simultaneously via any participating network subscriber.

FCC's 2005 Internet Policy Statement in using software applications to block or delay subscriber P2P file transfers.[76] The FCC chided Comcast for unduly interfering with Internet users' right to access the lawful Internet content and to use the applications of their choice:

> Although Comcast asserts that its conduct is necessary to ease network congestion, we conclude that the company's discriminatory and arbitrary practice unduly squelches the dynamic benefits of an open and accessible Internet and does not constitute reasonable network management. Moreover, Comcast's failure to disclose the company's practice to its customers has compounded the harm.[77]

Specifically, the FCC found that Comcast had deployed deep packet inspection[78] equipment throughout its network to monitor the content of its customers' Internet connections and to block specific types of P2P connections such as the use of BitTorrent software.[79]

Comcast used software that effectively substituted the company as the recipient of a customer's P2P file transfer session and enabled Comcast to issue a command to stop sending traffic. Comcast did this by forging so-called TCP reset packets[80] even though the company could have handled the current traffic volume without tampering with anyone's network use.[81] While Comcast initially disclaimed any responsibility for its customers' problems, tests conducted by the Associated Press and Electronic Frontier

---

76. *In re* Formal Complaint of Free Press & Public Knowledge Against Comcast Corp. for Secretly Degrading Peer-to-Peer Applications, 23 F.C.C.R. 13028 (2008) [hereinafter Comcast Internet Order].

77. *Id*. at 13028.

78. Deep packet is a deft technology that can "identify packets in real-time" and help a service provider to target and choke off peer-to-peer traffic. Kevin Werbach, *Breaking the Ice: Rethinking Telecommunications Law for the Digital Age*, 4 J. TELECOMM. & HIGH TECH. L. 59 (2005); *see also* Rob Frieden, *Internet Packet Sniffing and Its Impact on the Network Neutrality Debate and the Balance of Power Between Intellectual Property Creators and Consumers*, 18 FORDHAM INTELL. PROP. MEDIA & ENT. L. J., No. 3. 633-675 (2008).

79. The FCC noted the intrusiveness of Comcast's tactics in policing peer-to-peer traffic: "Comcast opens its customers' mail because it wants to deliver mail not based on the address or type of stamp on the envelope but on the type of letter contained therein." *Comcast Internet Order*, 23 F.C.C.R. at 13051 (citations omitted).

80. *Id*. ¶ 41, at 13051 (citations omitted) (describing the function and workings of Transmission Control Protocol, or TCP).

81. *Id*. at 13056. The FCC faulted Comcast's practice as "overinclusive" for interfering with network traffic without regard for actual evidence of congestion, times of peak usage, and types of disfavored applications. *Id*

Foundation[82] suggested that the company selectively interfered with subscribers' file-sharing sessions through P2P applications, a user practice that did not violate FCC rules or Comcast's service terms and conditions.[83] Comcast later admitted of targeting subscribers' P2P traffic for interference[84] but claimed that such interference was limited to times of network congestion.[85]

The FCC concluded that Comcast's practices did not constitute legitimate network management, because Comcast discriminated against specific applications without regard to whether the specific use actually caused congestion by depleting the Comcast network of sufficient bandwidth:

> On its face, Comcast's interference with peer-to-peer protocols appears to contravene the federal policy of "promot[ing] the continued development of the Internet" because that interference impedes consumers from "run[ning] applications . . . of their choice," rather than those favored by Comcast, and that interference limits consumers' ability "to access the lawful Internet content of their choice," including the video programming made available by vendors like Vuze. Comcast's selective interference also appears to discourage the "development of technologies"—such as peer-to-peer technologies—that "maximize user control over what information is received by individuals . . . who use the Internet" because that interference (again) impedes consumers from "run[ning] applications . . . of their choice," rather than those favored by Comcast.[86]

The FCC noted that Comcast had an anticompetitive motive to interfere with customers' use of P2P applications, because such software provides Internet users with a high-quality video access alternative to cable television services.[87] Such video distribution might compete with Comcast's video-on-demand ("VOD") service.[88] The FCC thus concluded that Comcast's prac-

---

82. *Id.* at 13031 (recounting the Associated Press's investigations and findings against Comcast).

83. *Id*. at 13032.

84. *Id*. at 13031 ("Following these tests, Comcast changed its account and admitted that it targets peer-to-peer traffic for interference.").

85. *Id*.

86. *Id*. at 13052 (citations omitted).

87. *Id.* at 13030

88. *Id.* After receiving Comcast's explanations for how it will treatment subscriber traffic in a nondiscriminatory manner, the FCC expressed concerns that Comcast's network management tactics might favor the company's telephone service offering vis-à-vis Internet telephone services of unaffiliated ventures. *See* Letter from Dana R. Shaffer, Chief, Wireline Competition Bureau, FCC & Matthew Berry, General Counsel, FCC, to Kathryn A.

tices were not minimally intrusive, as the company claimed, but rather invasive in substantially impeding subscribers' ability to access preferred content and applications.[89]

The FCC also concluded that Comcast exacerbated the situation by failing to disclose its practices to consumers.[90] Because Comcast hid the fact that it interfered with customers' use of P2P applications, customers had no way of knowing who was doing what to their connections.[91] As a result, the FCC found that many consumers experiencing difficulty using only certain applications would blame not Comcast, as it deserved, but rather on the applications themselves, to the detriment of the companies offering such applications in a competitive marketplace.[92] A majority of the FCC Commissioners concluded that although Internet access constitutes a lightly regulated information service, Comcast's discriminatory practices so harmed consumers and competitors alike that the FCC had to curb such behavior. The FCC's pro-consumer zeal and surprise intervention contrasts with its laissez-faire attitude toward perhaps even more egregious limitations wireless carriers impose on subscribers' access to content and software applications.[93]

Zachem, Vice President, Regulatory Affairs, Comcast Corp. (Jan. 18, 2009), *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-288047A1.pdf [hereinafter Comcast VoIP Letter]. The Commission's letter to Comcast seeks an explanation for the disparate treatment of VoIP services, particularly in light of Comcast's assertion that its VoIP service is "facilities-based." The letter appears to infer that facilities-based means that Comcast physically partitions its data bandwidth, thereby creating for its VoIP service stand alone links. For Comcast's response, see Letter from Kathryn A. Zachem, Vice President, Regulatory & State Legislative Affairs, Comcast Corp., to Dana Shaffer, Chief, Wireline Competition Bureau, FCC & Matthew Berry, General Counsel, FCC (Jan. 30, 2009), *available at* http://gullfoss2.fcc.gov/prod/ecfs/retrieve.cgi?native_or_pdf=pdf&id_document=6520194593. Comcast reported that a subscriber would experience a noticeable deterioration in service, including VoIP service not provided by Comcast, whenever a subscriber uses seventy percent of his or her "provisioned bandwidth" for fifteen minutes or more. Because Comcast separately provisions its VoIP service, congestion from cable modems would not affect their VoIP customers.

89. *Comcast Internet Order*, 23 F.C.C.R. at 13051-53.

90. Comcast's failure to disclose the company's practice to its customers has compounded the harm. *Id.* at 13051.

91. *Id.* at 13058-59.

92. *Id.* at 13058.

93. *See* United States Senate, Committee on Commerce, Science, and Transportation, Hearings, The Consumer Wireless Experience (June 17, 2009), http://commerce.senate.gov/public/index.cfm?FuseAction=Hearings.Hearing&Hearing_ID=03b81ffd-ba9f-42e6-8331-7c28f6d112b0.

B. Establishing FCC Jurisdiction over Internet-based Practices

In establishing its 2005 Internet Policy Statement, the FCC made clear its intention to incorporate the Policy Statement's principles "into its ongoing policymaking activities."[94] Along with the 2005 Internet Policy Statement, the FCC released its Wireline Broadband Order that largely eschewed regulation, but warned: "Should we see evidence that providers of telecommunications for Internet access or IP-enabled services are violating these principles, we will not hesitate to take action to address that conduct."[95]

Perhaps anticipating that Comcast would appeal its Order, the FCC extensively outlined its statutory authority for having substantive jurisdiction over Comcast's Internet-based practices and for reaching an administrative decision without a formal rulemaking session before adjudication. The FCC based its jurisdiction over disputes regarding discriminatory network management practices primarily on two statutory mandates: (1) section 230(b) of the Communications Act of 1934, as amended, where Congress stated the United States' policy "to preserve the vibrant and competitive free market that presently exists for the Internet" as well as "to promote the continued development of the Internet" and (2) section 706(a) of the Act, where Congress directed the FCC to "encourage the deployment on a reasonable and timely basis of advanced telecommunications capability to all Americans."[96] Comcast and other parties consider these broad statutory mandates too vague to justify the requirements the FCC decided to impose on Comcast.[97]

In particular, the FCC opted to exercise jurisdiction over P2P Internet connections based on ISPs' use of "communication by wire."[98] With that direct link to the general jurisdiction conferred under Title I of the Communications Act, the FCC exercised its ancillary jurisdiction on the premise that Comcast's practices undercut national Internet policy.

The FCC rejected Comcast's argument that any regulation of network access and management violated the FCC's 27-year old policy of leaving information services unregulated:

---

94. *Id*. at 13034 (quoting 2005 Internet Policy Statement, 20 F.C.C.R. 14986, ¶ 5 (2005)).

95. *In re* Appropriate Framework for Broadband Access to the Internet Over Wireline Facilities, 20 F.C.C.R. 14853 (2005).

96. The Commission also invoked the following statutory provisions as further justification for its decision to assume jurisdiction and to order a remedy: Communications Act of 1934, 47 U.S.C. §§ 151, 201, 256, 257, 601(a) (2006).

97. *Comcast Internet Order*, 23 F.C.C.R. at 13035, 13048.

98. 47 U.S.C. § 152(a) (2006).

> [T]he FCC previously indicated that it would not hesitate to take action in the event that providers violated the principles set forth in the *Internet Policy Statement*. Moreover, the FCC repeatedly has stated its willingness to exercise the full range of its statutory authority to ensure that providers of cable modem service meet the public interest in a vibrant, competitive market for Internet-related services. For instance, in the *Wireline Broadband Order*, the FCC found that it had jurisdiction over providers of broadband Internet access services and stated that "we will not hesitate to adopt any non-economic regulatory obligations that are necessary to ensure consumer protection and network security and reliability in this dynamically changing broadband era." Specifically with regard to cable modem service, in the 2002 *Cable Modem Declaratory Ruling* sustained by the Supreme Court in *Brand X*, the FCC sought comment on a wide range of statutory bases for exercising ancillary jurisdiction over cable modem service, including section 230(b) of the Act. The FCC also explicitly mentioned the blocking or impairing of subscriber access by a cable modem service provider as possible triggers for FCC "intervention."[99]

While deciding to act on a case-by-case basis, the FCC held that it did not have to conduct a rulemaking or hearing to investigate and remedy Comcast's practices:

> [T]o the extent that Comcast implies that our ancillary authority does not extend to adjudications but rather must first be exercised in a rulemaking proceeding, it is simply wrong. The question of whether the FCC has jurisdiction to decide an issue is entirely separate from the question of how the FCC chooses to address that issue. Perhaps more to the point, the D.C. Circuit has affirmed the FCC's exercise of ancillary authority in an adjudicatory proceeding and in the absence of regulations before.[100]

---

99. *Comcast Internet Order*, 23 F.C.C.R. at 13049-50 (citations omitted).

100. *Id.* at 13048-49 (citing CBS, Inc. v. FCC, 629 F.2d 1, 26–27 (1980) (reasoning that the Commission had, in the context of an adjudication, reasonably construed its ancillary authority to encompass television networks), *aff'd*, 453 U.S. 367 (1981)); *In re* Complaint of Carter-Mondale Presidential Comm., Inc. against the ABC, CBS & NBC Television Networks, 74 F.C.C.2d 631, ¶ 25 n.9 (1979) ("Our power to adjudicate complaints involving requests for access to the networks is surely 'reasonably ancillary to the effective performance of the Commission's various responsibilities.' ") (quoting United States v. Sw. Cable Co., 392 U.S. 157, 178 (1968)); *see also* Negrete-Rodriguez v. Mukasey, 518 F.3d 497, 503-04 (7th Cir. 2008) ("An agency is not precluded from announcing new principles in an adjudicative proceeding rather than through notice-and-comment rule-making."); New York State Comm'n on Cable Television v. FCC, 749 F.2d 804, 815 (D.C. Cir. 1984)

C. Remedying Breaches of the Internet Nondiscrimination Policy

The FCC outlined several network management tactics that Comcast could lawfully impose on subscribers,[101] including capping average users' capacity and then charging overage fees, and throttling back the connection speeds of high-capacity users (rather than any user who relies on P2P technology, however infrequently.[102]

Although Comcast and other ISPs had pursued each of these options, the FCC determined that Comcast had to revise its network management practices and operate more transparently.[103] Within thirty days of release of the FCC's Order, Comcast had to:

> Disclose the details of its discriminatory network management practices to the FCC;
>
> Submit a compliance plan for eliminating these discriminatory management practices by the end of the year; and
>
> Disclose to customers and the FCC the new network management practices.[104]

The FCC warned Comcast that non-compliance with the steps set forth in the Order would lead to immediate interim injunctive relief requiring the company to suspend its discriminatory network management practices pending a hearing.[105]

For cable television ventures, the FCC had grave concerns about their potential for engaging in discriminatory practices when providing Internet access to subscribers. For wireless carriers, however, it is inconsistent and suspect for the FCC not to have the same concerns about similar tactics. While consumers may have only one cable television venture available in

---

(upholding adjudicatory decision that preempted certain state and local satellite television regulations under Commission's ancillary authority).

101. *Comcast Internet Order*, 23 F.C.C.R. at 13057-58.

102. *Id.*

103. Comcast's claim that it has always disclosed its network management practices to its customers is simply untrue. Although Comcast's Terms of Use statement may have specified that its broadband Internet access service was subject to "speed and upstream and downstream rate limitations," such vague terms are of no practical utility to the average customer. *Id*. at 13059-60 ("Our overriding aim here is to end Comcast's use of unreasonable network management practices, and our remedy sends the unmistakable message that Comcast's conduct must stop.").

104. *Id*. at 13059-60

105. *Id*. at 13060.

any particular locality, the FCC notes that alternative broadband access providers exist, ostensibly providing equivalent or more competition than that provided by wireless carriers.[106] Although users can change broadband providers in an ostensibly competitive marketplace, the FCC was not content to allow market forces to punish Comcast through customer churn and migration to other carriers. Arguably, wireless carriers' discriminatory practices are more egregious in restricting, burdening, and thwarting subscribers' lawful access to content and software applications. Remarkably, even though service agreements have limited subscribers' choice to switch carriers, the FCC apparently sees no need to bind wireless carriers to the Internet Freedoms or to nondiscrimination requirements of traditional common carriers.

VII. ACHIEVING A LEVEL PLAYING FIELD AND SERVING THE PUBLIC INTEREST

Industry observers and interested parties disagree over whether wireless carriers robustly compete, or whether four national carriers share an oligopoly with over ninety percent of combined market shares. The FCC believes in the former and repeatedly eschews regulation based on market discipline competition provides.[107] Remarkably the FCC refuses to believe that the competition cable television operators face from satellite and new wireline ventures, including Verizon and AT&T, will foreclose anticompetitive practices. The FCC has intervened to remedy anticompetitive and anti-consumer restrictions on Internet access imposed by Comcast, a largely unregulated information service provider,[108] but appears willing to tolerate similar forms of blocked or restricted Internet access imposed by wireless carriers.

The FCC has articulated no legal basis for creating such lack of parity in regulation, and in its refusal to remedy anticompetitive abuses by wireless carriers. In the United States, wireless carriers have actively and aggressive-

---

106. *See In re* Dev. of Nationwide Broadband Data to Evaluate Reasonable & Timely Deployment of Advanced Servs. to All Ams., Improvement of Wireless Broadband Subscribership Data, & Dev. of Data on Interconnected Voice over Internet Protocol (VoIP) Subscribership, 23 F.C.C.R. 9691 (2008); *In re* Inquiry Concerning the Deployment of Advanced Telecomms. Capability to All Ams. in a Reasonable & Timely Fashion, & Possible Steps to Accelerate Such Deployment Pursuant to Section 706 of the Telecomms. Act of 1996, 23 F.C.C.R. 9615, 9616 (2008); WIRELINE COMPETITION BUREAU, FCC, HIGH-SPEED SERVICES FOR INTERNET ACCESS: STATUS AS OF JUNE 30, 2007 (2008), *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-280906A1.doc.

107. *In re* Implementation of Section 6002(b) of the Omnibus Budget Reconciliation Act of 1993, 23 F.C.C.R. 2241 (2008).

108. *Comcast Internet Order*, 23 F.C.C.R. at 13057-58.

ly limited subscribers' handset flexibility and freedom with impunity. Nonetheless, the FCC has been silent on regulatory requirements for wireless carriers that provide both content and conduit for next generation ICE. Disposed to supporting marketplace competition and deregulation, the FCC has given wireless carriers carte blanche freedom to operate as though they had no regulatory burdens even for still provided conventional telecommunications services.

The FCC's unwillingness to apply consumer-protection statutes lead to uneven regulation and circumscribe consumer freedom in handset use. Rather than wait for a consumer revolt, the FCC should state that:

> (1) Wireless subscribers have the right to attach any handset that complies with standards designed to protect networks from technical harm; wireless operators should bear the burden of proving that a particular handset would cause technical harm and therefore should not receive FCC certification;
>
> (2) Wireless subscribers have the right to use their handsets to access any service, software, application and content available by subscriber imputed commands or instructions. The FCC should expressly state that wireless operators should have an affirmative duty to receive, switch, route and transmit such subscriber keyed commands or instructions; and
>
> (3) Suppliers of software, applications, services and content accessible via wireless networks have the right to offer them to subscribers subject to a reasonable determination by carriers that such access will not cause technical harm to the carriers' networks. The FCC should reserve the right to mediate and resolve disputes over technical compatibility of any software, applications, services and content accessible via a wireless carrier network.

Additionally the FCC needs to watch out for instances of excessive market consolidation. When wireless carriers acquire market share by buying out competitors, consumers have fewer opportunities to "vote with their feet" against their existing carrier. Already, wireless carriers can engage in "consciously parallel" conduct where they offer similar terms and conditions for service. For example, no carrier offers in the U.S. a discounted rate for existing or prospective customers who already have handsets and do not need subsidized new phones. Indeed, most cellular carriers advertise how well their networks operate, without much discussion about price and service versatility.

Lastly, the FCC needs to force wireless carriers to operate in a more transparent and forthright manner so that subscribers understand the total cost of service and the limitations imposed on their handsets. Many wireless carriers have managed to insert into consumers' bills surcharges that appear as a tax or other type of compulsory fee. Without the duty to file tariffs and standard service terms and conditions, some wireless carriers can impose unanticipated charges. When handsets can access a much larger array of ICE services, it becomes essential that wireless carriers fully disclose any offer of superior access to content and preferential access for specific content providers. Not all quality-of-service and price differentials violate a reasonable conceptualization of fairness and neutrality. On the other hand, failure to disclose discriminatory pricing and service is improper and anti-competitive.

Requiring transparency and fair dealing by carriers providing the underlying transmission capacity for ICE services serves the national interest. Unfortunately, the issue of imposing such straightforward requirements has become submerged in a larger debate about how much flexibility ISPs should have to diversify services and what regulator-created incentives are necessary to encourage investment in next generation networks. The so-called network neutrality debate[109] addresses what constitutes lawful price

---

109. *See* Rob Frieden, *Internet 3.0: Identifying Problems and Solutions to the Network Neutrality Debate*, 1 INT'L J. COMM. 461 (2007); Rob Frieden, *Network Neutrality or Bias? Handicapping the Odds for a Tiered and Branded Internet*, 29 HASTINGS COMM. & ENT. L.J. 171, 171-216 (2007); *see also* Barbara A. Cherry, *Using Network Neutrality to Eliminate Common Carriage Threatens Free Speech and the Postal System*, 33 N. KY. L. REV. 483 (2006); Brett M. Frischmann & Barbara van Schewick, *Network Neutrality and the Economics of an Information Superhighway: A Reply to Professor Yoo*, 47 JURIMETRICS J. 383 (2007); Bill D. Herman, *Opening Bottlenecks: On Behalf of Mandated Network Neutrality*, 59 FED. COMM. L.J. 103 (2006); Mark A. Lemley & Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 UCLA L. REV. 925 (2001); J. Gregory Sidak, *A Consumer-Welfare Approach to Network Neutrality Regulation of the Internet*, 2 J. COMPETITION L. & ECON. 349 (2006); Adam Thierer, *Are "Dumb Pipe" Mandates Smart Public Policy? Vertical Integration, Net Neutrality, and the Network Layers Model*, 3 J. TELECOMM. & HIGH TECH. L. 275 (2005); Barbara van Schewick, *Towards an Economic Framework for Network Neutrality Regulation*, 5 J. TELECOMM. & HIGH TECH. L. 329 (2007); Tim Wu, *Network Neutrality, Broadband Discrimination*, 2 J. TELECOMM. & HIGH TECH. L. 141 (2003); Christopher S. Yoo, *Network Neutrality and the Economics of Congestion*, 94 GEO. L.J. 1847 (2006); Christopher S. Yoo, *Would Mandating Broadband Network Neutrality Help or Hurt Competition? A Comment on the End-to-End Debate*, 3 J. TELECOMM. & HIGH TECH. L. 23 (2004); Craig McTaggart, *Was The Internet Ever Neutral?* (Sept. 30, 2006), *available at* http://web.si.umich.edu/tprc/papers/2006/593/mctaggart-tprc06rev.pdf (presented at the 34th Research Conference on Communication, Information and Internet Policy, George Mason University School of Law).

and service discrimination by ISPs, but opponents have managed to frame the debate as an unlawful attempt to impose common carrier responsibilities. While ISPs do not operate as common carriers, wireless telecommunications service providers do, in addition to their information and video programming, delivery services. The FCC faces a regulatory quandary when wireless carriers augment telecommunications with new ICE services. The carriers claim that the FCC no longer has any legal basis for regulating any wireless service, because the FCC cannot subject a single enterprise to more than one regulatory regime and because competition will prevent any anticompetitive practices. However, the onset of new services, which may qualify as something other than a common carriage telecommunications service, does not by itself vitiate or eliminate the initial regulatory requirements. When handsets offer an amalgam of telecommunications, information, and video services, the FCC and wireless carriers will need to learn how to operate in an environment where two or more regulatory regimes apply.

COPYRIGHT INFRINGEMENT IN THE INTERNET AGE—PRIMETIME FOR HARMONIZED CONFLICT-OF-LAWS RULES?

By Anita B. Frohlich†

TABLE OF CONTENTS

---

† National University San Diego; LL.M. in Comparative Law (University of San Diego School of Law); Master's of Law (Université Aix-Marseille III School of Law, France); JD (University of Passau School of Law, Germany). The author would like to express her gratitude to Professor Lisa Ramsey, University of San Diego School of Law, for her helpful suggestions and comments, and to the staff of the Berkeley Technology Law Journal for excellent editing. The author maintains a weblog at www.comparelex.org.

I. INTRODUCTION

The staggering quantity of copyright infringement on the Internet has exposed a fundamental limitation of current international copyright law. A single act of unauthorized uploading of copyrighted material can result in copyright infringements by numerous Internet users in various countries. The following hypothetical shows the legal difficulties that are raised in such situations:

> Bill, a United States citizen who resides in Great Britain, uploads digitalized copies of the work of Françoise from his residence to a server in Great Britain. Françoise, a French citizen who lives in Belgium, has published her work solely in France and holds a French copyright. Bill has not obtained permission to publish Françoise's work. The infringing material is then downloaded in Germany, the United States and Great Britain. Françoise sues Bill in Belgium for damages.[1]

Determining which law should apply in this situation is a difficult question. A court could apply French law because Françoise is a French citizen holding a French copyright for work published in France, Belgian

1. Inspired by Alexander Peinze, Internationales Urheberrecht in Deutschland und England 119 (2002).

law as Françoise lives in Belgium and sues Bill in a Belgian court, British law as the law of Bill's country of residence and of the place where Françoise's work was both uploaded without permission and subsequently downloaded by Internet users, United States' law because Bill is a U.S. citizen and the infringing work was downloaded in the United States, or even German law as the work was also downloaded there.

International copyright law does not provide a satisfactory resolution for such an international case since the main pillar of intellectual property law—the principal of territoriality—implies nationally limited application of copyright law. Accordingly, the exclusive rights afforded to a copyright owner can only be exercised within the borders of a given country.[2] This national limitation of copyrights is in pronounced contrast to the universal validity of other rights (e.g., contractual rights).[3]

In order to decide the above-described case, one must consider not only copyright law, but also conflict-of-laws methods. Traditionally, however, there has been little exchange between conflict-of-laws and intellectual property scholars.[4] This mutual ignorance stems from the fact that intellectual property law used to reflect the prototypical expression of sovereign or national interests and therefore involves the exclusive application of domestic law. As a result of the growing distribution of copyrighted work over the Internet, however, intellectual property scholars must concern themselves increasingly with conflict-of-laws issues. This Article establishes some common ground for the previously distinct fields of conflict-of-laws and international intellectual property law. Specifically, the Article demonstrates how to apply conflict-of-laws theory to international copyright cases while making allowances for the special character of international intellectual property law.

As the Bill-Françoise hypothetical above shows, there are many laws that could possibly apply to cases of internet copyright infringement. In this context, the following choice-of-law rules are available: *lex fori*, *lex loci delicti*, and *lex protectionis* rule.

The *lex fori* rule postulates application of the law of the country where the court deciding the infringement case is situated (law of the forum).[5]

---

2. *See* Richard Fentiman, *Choice of Law and Intellectual Property*, *in* INTELLECTUAL PROPERTY AND PRIVATE INTERNATIONAL LAW: HEADING FOR THE FUTURE 129, 137 (Josef Drexl & Annette Kur eds., 2005).

3. *See* JAN KROPHOLLER, INTERNATIONALES PRIVATRECHT 535 (2004); Fentiman, *supra* note 2, at 138.

4. *See* Graeme B. Dinwoodie, *International Intellectual Property Litigation: A Vehicle for Resurgent Comparativist Thought?*, 49 AM. J. COMP. L. 429, 429 (2001).

5. BLACK'S LAW DICTIONARY 760 (abr. 8th ed. 2005).

Application of this rule inevitably leads to the court applying its own law.[6] In contrast, the *lex loci delicti* rule provides for application of the law of the country where the infringing conduct has occurred (place of the wrong).[7] Application of the *lex loci delicti* may require a court to apply foreign law in cases where courts assume international jurisdiction with regard to infringements of foreign rights.[8] Finally, the *lex protectionis* rule applies the law of the country for which protection is sought.[9] Contrary to the *lex fori* rule, the *lex protectionis* rule may give rise to the application of a foreign law.[10] Although the *lex loci delicti* and the *lex protectionis* rules often result in the application of the same law, the *lex protectionis* rule generally tends to be broader, governing many different aspects of copyright law, not necessarily restricted to issues of infringement.[11]

In the Bill-Françoise hypothetical above, Françoise filed suit in a Belgian court. If the Belgian court applied the *lex fori* rule, the court would apply Belgian law to the dispute. If the court applied the *lex loci delicti* rule, it would have to apply foreign law since none of the downloads occurred in Belgium. Complicating matters further, application of the *lex loci delicti* may require application of laws from multiple countries if the court considered the location of the wrong to be the country where the infringing act had its effect (i.e., where the downloads actually occurred). Under this approach, the court in the Bill-Françoise hypothetical would have to apply German, British, and United States law. Since the Bill-Françoise hypothetical focuses on the infringement aspect of copyright violation, the *lex protectionis* rule would give rise to the same outcome as the *lex loci delicti*.

This plethora of approaches to choice of law in copyright infringement cases still inadequately addresses the challenges posed by the ease of content-distribution through the Internet. In response, multiple possible solu-

---

6. Choice-of-law rules may refer to application of a country's law either including or excluding the choice-of-law rules of this country. If choice-of-law rules are included in the referral, a mechanism of referring back and forth (renvoi) may ensue. For simplification reasons, this Article does not consider whether a particular choice-of-law rule includes a referral to choice-of-law rules or not. Each time the Article states that a choice-of-law rule leads to application of a country's law, either application of the country's substantive law or application of the country's choice-of-law rules can result.

7. *See* BLACK'S LAW DICTIONARY, *supra* note 5, at 760.

8. *See* EUGEN ULMER, INTELLECTUAL PROPERTY RIGHTS AND THE CONFLICT OF LAWS 11 (1978).

9. *See, e.g.*, MIREILLE VAN EECHOUD, CHOICE OF LAW IN COPYRIGHT AND RELATED RIGHTS 105 (2003).

10. *See* ULMER, *supra* note 8, at 11.

11. *See* VAN EECHOUD, *supra* note 9, at 105.

tions have recently emerged on the international plane. In the European Union, a new "Regulation of the European Parliament and of the Council on the law applicable to non-contractual obligations" (Rome II) now includes a provision relating to the infringement of intellectual property rights.[12] The American Law Institute has drafted its "Intellectual Property: Principles Governing Jurisdiction, Choice of Law, and Judgments in Transnational Disputes" (ALI Principles), a comprehensive regulation of conflict-of-laws issues specifically focused on intellectual property rights. The ALI Principles specifically include provisions relating to internet copyright infringement.[13]

The aim of this Article is to explore the best approach to the ever-increasing problem of determining which law to apply to cases of multi-national copyright infringement on the Internet. Some critics have tried to dismiss the importance of choice-of-law considerations in copyright, either by claiming that international intellectual property agreements already answer the questions[14] or by pointing to existing consensus in practice.[15] This Article will establish the flippancy of such arguments especially in the light of the incessant distribution of copyright infringing content over the Internet and will emphasize the importance of focusing on the intersection of conflict of laws and intellectual property law. For this purpose, Part II analyzes the status quo of conflict-of-laws rules in international copyright law. It will reveal a portfolio of approaches on the national, regional, and international plane and by using the above-introduced hypothetical illustrate the short-comings of the current legal situation, particularly with regard to multinational internet copyright infringement. Part III will then examine recent and still pending harmonization efforts in the area of choice of law and intellectual property law, evaluate their value for resolving choice-of-law issues in multi-national internet copyright infringement cases, and pave the way for the successful cross-fertilization of these two previously distinct fields.

---

12. *See* Council Regulation 864/2007, art. 8, 2007 O.J. (L 199) 40 (EC) [hereinafter Rome II].

13. *See* AMERICAN LAW INSTITUTE [ALI], INTELLECTUAL PROPERTY: PRINCIPLES GOVERNING JURISDICTION, CHOICE OF LAW, AND JUDGMENTS IN TRANSNATIONAL DISPUTES: PROPOSED FINAL DRAFT (2007).

14. *See* Dinwoodie, *supra* note 4.

15. *See* Josef Drexl, *The Proposed Rome II Regulation: European Choice of Law in the Field of Intellectual Property*, *in* INTELLECTUAL PROPERTY AND PRIVATE INTERNATIONAL LAW: HEADING FOR THE FUTURE 151, 152 (Josef Drexl & Annette Kur eds., 2005).

II. CHOICE OF LAW AND COPYRIGHT INFRINGEMENT: A PORTFOLIO OF APPROACHES

Harmonization and streamlining efforts aside, United States and European courts tend to approach the issue of conflict of laws in copyright law differently. Even within Europe, there exist remarkable differences. This part of the Article presents the current portfolio of different approaches within the United States and Europe.

A. The Situation in the United States

In the United States, copyright of published and unpublished works is strictly protected by federal law.[16] However, the United States did not join the Berne Convention for the Protection of Literary and Artistic Works,[17] an international instrument establishing minimum standards for copyright protection, until 1989.[18] This long hesitation was due to the fact that Congress considered U.S. copyright law as sacrosanct, taboo for regulation by international law,[19] an attitude which is also reflected in United States jurisprudence. At the beginning of the 20th century, the United States Supreme Court determined that copyright law was strictly territorial.[20] As a consequence, common law courts were inclined to invoke *forum non conveniens* when foreign copyright law was involved.[21] The accession of the United States to the Berne Convention initiated a gradual reconsideration of international copyright law while contemporaneously creating new challenges when foreign copyright law was involved. The Ninth and Second Circuits have taken conflicting approaches to address these challenges. The Ninth Circuit viewed the Berne Convention as requiring a *lex fori* approach, while the Second Circuit ultimately ignored the Berne Con-

16. *See* U.S. CONST. art. 1, § 8, cl. 8.

17. *See* Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as last revised at Paris, July 24, 1971, S. Treaty Doc. No. 99-27, 1161 U.N.T.S. 3 (1986) [hereinafter Berne Convention].

18. *See* WIPO Administered Treaties: Berne Convention Contracting Parties: Details on United States of America, http://www.wipo.int/treaties/en/Remarks.jsp? cnty_id =1045C (last visited May 6, 2008).

19. *See* David E. Miller, *Finding A Conflicts Issue in International Copyright Litigation: Did the Second Circuit Misinterpret the Berne Convention in Itar-Tass?*, 8 CARDOZO J. INT'L & COMP. L. 239, 244 (2000).

20. *See, e.g.*, United Dictionaries Co. v. G. & C. Merriam Co., 208 U.S. 260 (1908) (holding that the American copyright statute does not require notice of the American copyright on books published abroad and sold only for use there).

21. *See, e.g.*, Subafilms, Ltd. v. MGM-Pathé Commc'n Co., 24 F.3d 1088, 1095 (9th Cir. 1994) (en banc).

vention and used a broader interests approach to choose the proper law on a case-by-case basis.

1. *The Ninth Circuit Approach*

In *Subafilms Ltd. v. MGM-Pathe Communications Co.*, the United States Court of Appeals for the Ninth Circuit considered whether United States copyright law applied where the authorization of acts that would lead to copyright infringements abroad occurred in the United States.[22] The court held en banc that the U.S. Copyright Act did not apply to cases of secondary infringement where the authorization of infringing acts occurred abroad, even if the acts constituted infringement if committed in the United States.[23] The *Subafilms* decision is thus characterized by a strict understanding of the principle of territoriality.[24] The Ninth Circuit based its strict interpretation of territoriality on the importance of comity in international copyright relations. In particular, the court emphasized that the extraterritorial application of U.S. copyright law impaired comity because it imposed application of United States copyright law on issues that would otherwise fall within the sovereignty of another state.[25] Accordingly, the Ninth Circuit declined to decide the case on the grounds of *forum non conveniens*.

Such a restrictive approach to conflict-of-laws issues in copyright law will often prove detrimental for the copyright holder as it may bar him from bringing a lawsuit in a country where the copyright infringer has his assets. For example, if in the Bill-Françoise hypothetical, all the courts involved adopted an interpretation similar to the Ninth Circuit, Françoise would not be able to sue Bill in either his home country (United States) or his country of residence (Great Britain). Rather, Françoise would be limited to seeking redress in a French court that would then apply its own law. If Françoise won her lawsuit, she would have to enforce the French judgment in order to collect damages. And if Bill had no assets in France, such enforcement would be complex and possibly warrant further judicial proceedings.[26]

From a conflict-of-laws perspective, the reasoning of the Ninth Circuit in *Subafilms* is flawed because it combines questions of jurisdiction and

---

22. *Id.*
23. *Id.* at 1095.
24. *Id.* at 1095-96.
25. *Id.* at 1097.
26. For the enforcement of judgments, see, for example, U.S. Dept. of State, Enforcement of Judgments, http://travel.state.gov/law/info/judicial/judicial_691.html (last visited Feb. 27, 2009).

choice of law. The court based its denial of the claim on the inappropriateness of applying U.S. copyright law extraterritorially.[27] This argument already implies that the assumption of subject-matter jurisdiction would inevitably result in application of *lex fori*. Such an approach ignores the two-step analysis fundamental to conflict-of-laws cases, requiring that the questions of jurisdiction and choice of law be decided subsequently.[28] If a court has jurisdiction, this does not mean that forum law automatically applies to the case: the determination of the applicable law is another separate issue from jurisdiction.[29]

Consequently, by asserting *forum non conveniens* due to the lack of extraterritorial application of U.S. copyright law without first determining whether such law would be applicable in the first place, *Subafilms* lacks a clear-cut distinction between jurisdiction and choice of law and, as a result, is inconsistent with conflict-of-laws standards. In other words,

> to say that each country is authorised to legislate its own copyright, and that, therefore, it cannot by definition be applied beyond its borders, is to negate the existence of private international law, or less drastically: to reduce it to the maxim that all courts should always apply their own law.[30]

In addition, the Ninth Circuit appears to suggest that article 5(1) of the Berne Convention—which postulates the principle of national treatment[31]—points to a choice-of-law rule.[32] In *Subafilms*, the court states that "[a]lthough the treaties do not expressly discuss choice-of-law issues, it is commonly acknowledged that the national treatment principle *implicates* a rule of territoriality."[33] This statement indicates that the court considered the national treatment principle to have choice-of-law implications leading to application of the *lex fori* as the most territorial choice-of-law rule.[34]

---

27. *Subafilms*, 24 F.3d at 1097-98.

28. *See* EUGENE SCOLES ET AL., CONFLICT OF LAWS (3d ed. 2000).

29. *Id.*

30. *See* VAN EECHOUD, *supra* note 9, at 97.

31. The national treatment principle stipulates that a country shall treat authors from other member states of the Berne Convention the same as its domestic authors. *See, e.g.*, WIPO, Summary of the Berne Convention for the Protection of Literary and Artistic Works (1886), http://www.wipo.int/treaties/en/ip/berne/summary_berne.html.

32. *See* Graeme W. Austin, *Importing* Kazaa *Exporting* Grokster, 22 SANTA CLARA COMPUTER & HIGH TECH. L.J. 577, 595 n.76 (2006).

33. Subafilms, Ltd. v. MGM-Pathé Commc'ns Co., 24 F.3d 1088, 1097 (9th Cir. 1994) (emphasis added) (citations omitted).

34. *See* Creative Tech. v. Aztech Sys. Pte, Ltd., 61 F.3d 696, 700 (9th Cir. 1995).

The character of the national treatment principle, however, has been subject to debate.[35] While case law partially supports the view that the national treatment principle constitutes a proper choice-of-law rule,[36] many scholars see the issue of national treatment as a precursor to the question of the applicable law.[37] On the other hand, another group of intellectual property scholars contend that the national treatment principle merely stipulates a principle of nondiscrimination and has limited choice-of-law implications or even none at all.[38] Opponents of the characterization of the national treatment principle as a choice-of-law standard get support from their conflict-of-laws colleagues.[39] According to general conflict-of-laws theory, choice-of-law rules determine, among other things, which *law* to apply.[40] The national treatment principle of the Berne Convention, however, stipulates that each member state should accord foreign authors "the *rights* which their respective laws do now or may hereafter grant to their nationals."[41] If this is true, then the national treatment principle would only come to play once a suitable choice-of-law rule has determined the applicable law in the first place.

Historical evidence also points away from considering the national treatment principle as a choice-of-law rule.[42] At the time that Berne was negotiated, copyright law was not expressly classified as public or private law, and a uniform approach to conflict of laws did not exist.[43] Furthermore, the Berne Convention was drafted in efforts to harmonize international copyright law, which strongly suggests that the drafters of the Berne Convention did not focus on conflict issues, since the whole study of conflict of laws is based on the diversity of laws, not their harmonization.[44]

There is also a logical argument against applying the national treatment principle as a choice-of-law rule. According to article 5(1) of the

---

35. *See* Drexl, *supra* note 15, at 165.

36. *See* Murray v. British Broad. Corp., 81 F.3d 287, 290 (2d Cir. 1996) (citing *Creative Tech.*, 61 F.3d at 700).

37. Drexl, *supra* note 15, at 165 n.64 (referring to Koumantos).

38. *See* PAUL GOLDSTEIN, INTERNATIONAL COPYRIGHT: PRINCIPLES, LAW, AND PRACTICE 89 (2001); VAN EECHOUD, *supra* note 9, at 107 (limited choice-of-law implications); William Patry, *Choice of Law and International Copyright*, 48 AM. J. COMP. L. 383, 405 (2000) (no choice-of-law implications).

39. *See* P.L.C. Torremans, *Intellectual Property and Private International Law: Heading for the Future*, 5 EUR. INTELL. PROP. REV. 312, 469 (2006).

40. *Id.*

41. Berne Convention, *supra* note 17, art. 5(1) (emphasis added).

42. *See* VAN EECHOUD, *supra* note 9, at 92.

43. *Id.* at 92-93.

44. *Id.* at 93.

Berne Convention, national treatment is only accorded to foreigners and not nationals of the country of origin.[45] National treatment, however, means that foreigners shall be treated equally to nationals.[46] For example, suppose that one country applies the *lex fori* rule for its nationals with regards to the law applicable to copyright infringement. An interpretation of the national treatment principle, however, would lead to application of *lex protectionis*. Provided that both conflict rules conform to the minimum standards set by the Berne Convention,[47] the above-described situation would result in unequal treatment of nationals and foreigners and thus conflict with the idea underlying the national treatment principle.[48] Thus, there are strong arguments against construing the national treatment principle as a choice-of-law rule, yet courts sometimes decide this question otherwise.[49]

Even if the national treatment principle is considered to be a choice-of-law rule, the exact nature of the rule is not yet established. This problem becomes clear in *Subafilms*. Although the Ninth Circuit states that "the national treatment principle implicates a rule of territoriality," it is unclear what "rule of territoriality" means for conflicts purposes.[50] Some argue that such a view leads to application of the *lex protectionis*.[51] Others find that territoriality inevitably implies the *lex loci delicti*, the generally accepted choice-of-law rule for torts.[52] A third interpretation of the national treatment principle would lead to the *lex fori*.[53] This ambiguity further supports the argument against characterizing the national treatment principle as a choice-of-law rule.

---

45. *See* Berne Convention, *supra* note 17, art. 5(1).

46. *Id.* ("Authors shall enjoy, in respect of works for which they are protected under this Convention, in countries of the Union other than the country of origin, *the rights which their respective laws do now or may hereafter grant to their nationals* . . . .") (emphasis added).

47. Works and rights to be protected and duration of protection. *See* Berne Convention, *supra* note 17, art. 7 (duration of protection), art. 8-9, 11-12, 14, 14ter (economic rights) and art. 6bis (moral rights). For a list of the minimum standards of protection, see also WIPO, Summary of the Berne Convention for the Protection of Literary and Artistic Works (1886), http://www.wipo.int/treaties/en/ip/berne/summary_berne.html (last visited Apr. 2, 2009).

48. *See* VAN EECHOUD, *supra* note 9, at 127.

49. *See, e.g.*, Subafilms, Ltd. v. MGM-Pathe Commc'ns Co., 24 F.3d 1088 (9th Cir. 1994).

50. *Cf.* VAN EECHOUD, *supra* note 9, at 96 (on the basis of Dutch case law).

51. *See* Drexl, *supra* note 15, at 166.

52. *See* Kaspar Spoendlin, *Der internationale Schutz des Urhebers*, 107 UFITA 11, 17 (1988).

53. *See Subafilms*, 24 F.3d at 1097.

In sum, *Subafilms* presents two major challenges from a conflict-of-laws perspective. First, it commingles questions of jurisdiction and choice of law by basing its *forum non conveniens* decision on the lack of extraterritorial application of U.S. copyright law. Second, it suggests that the national treatment principle of the Berne Convention implicates a choice-of-law rule for cases of international copyright infringement, in opposition to arguments against such an interpretation that are expressly contained within the Berne Convention. The following Section illustrates that critics of *Subafilms* are further supported by subsequent case law in the Second Circuit.

2. *The Second Circuit Approach*

Originally, the Second Circuit assumed an approach similar to *Subafilms*. In *Murray v. British Broadcasting Corp.*, the Second Circuit stated that national treatment was a choice-of-law rule.[54] However, contrary to the Ninth Circuit in *Subafilms*, the Second Circuit in *Murray* clearly acknowledged that interpreting the national treatment doctrine as a choice-of-law rule resulted in application of *lex loci delicti*.[55] The more definite approach by the Second Circuit was less ambiguous than the Ninth Circuit's decision in *Subafilms*.

The Second Circuit's interpretation of the national treatment principle changed remarkably with its decision in *Itar-Tass Russian News Agency v. Russian Kurier, Inc.*[56] *Itar-Tass* is considered one of the leading United States cases on choice-of-law issues in copyright law.[57] The case involved the unauthorized distribution, in New York, of newspaper articles that were originally published in Russia. In contrast to the Ninth Circuit in *Subafilms*, the Second Circuit in *Itar-Tass* extensively examined the issue of applicable law. Based on its assessment of the existing federal and state law, the court found no appropriate choice-of-law rule for copyright infringement cases and, therefore, created its own rules.[58] This implies that the court did not consider the national treatment principle to be a valid choice-of-law rule. Rather, the court found that the "principle of national treatment simply assure[d] that if the law of the country of infringement

54. *See* Murray v. British Broad. Corp., 81 F.3d 287, 290 (2d Cir. 1996).

55. *Id.* at 290, 293.

56. Itar-Tass Russian News Agency v. Russian Kurier, Inc., 153 F.3d 82 (2d Cir. 1998).

57. *See* Graeme Dinwoodie, *Resolution through Conflict-of-laws: Remarks by Graeme Dinwoodie*, 30 BROOK. J. INT'L L. 885, 890 (2005).

58. *Itar-Tass Russian News Agency*, 153 F.3d at 90.

applie[d] to the scope of substantive copyright protection, that law [would] be applied uniformly to foreign and domestic authors."[59]

The court noted that different legal issues warranted different choice-of-law rules.[60] Such *depeçage*, i.e. the separation of one comprehensive legal relationship to several legal issues to which different choice-of-law rules are applied, is not uncontested among American conflict-of-laws scholars.[61] The Second Circuit distinguished between issues of ownership and infringement for choice-of-law purposes. On the issue of copyright infringement, the court applied the *lex loci delicti*.[62] Interestingly, the court embedded its application of the *lex loci delicti* in a broader interest analysis.[63] The court noted the relevance of determining both the ownership of an interest and the nature of that interest, but concluded that, in the case at hand, application of U.S. law to the infringement issue was an obvious choice since the infringement occurred in the United States and the defendant was a U.S. corporation.[64]

The Second Circuit's choice of *lex loci delicti* through a broader interest approach is problematic for two reasons. First, a broader interest analysis can lead to legal indeterminacy or even to forum favoritism.[65] For instance, the Second Circuit in the *Itar-Tass* decision articulated that "United States law would still apply to infringement issues, since not only is this country the place of the tort, but also the defendant is a United States corporation."[66] Nonetheless, consideration of state interests in conflict of laws occurs quite frequently in United States courts.[67]

Second, the application of the *lex loci delicti* could lead to a multitude of applicable laws.[68] This difficulty is best illustrated by applying the *lex loci delicti* to the Bill-Françoise hypothetical. The *lex loci delicti* rule would apply the law of the place of infringement. Yet the location of in-

---

59. *Id.* at 89.

60. *Id.* at 90.

61. *See* SYMEON SYMEONIDES, THE AMERICAN CHOICE-OF-LAW REVOLUTION: PAST, PRESENT AND FUTURE 31-33 (2006).

62. *See Itar-Tass Russian News Agency*, 153 F.3d at 91. For the question of copyright ownership, the court determined the law of the country of origin of the copyright to be applicable in this specific case.

63. *Id.*

64. *Id.*

65. *See* Klaus Schurig, *Interessenjurisprudenz contra Interessenjurisprudenz im IPR —Anmerkungen zu Flessners Thesen*, 59 RABELS ZEITSCHRIFT FUER AUSLAENDISCHES UND INTERNATIONALES PRIVATRECHT 229, 242 (1995) (F.R.G.).

66. *Itar-Tass Russian News Agency*, 153 F.3d at 91.

67. *See* SYMEONIDES, *supra* note 61, at 373.

68. *See* Dinwoodie, *supra* note 4, at 440.

fringement is uncertain when Bill uploaded works that infringe on Françoise's copyright from his home computer in Great Britain to a server in Great Britain, and these works are then downloaded in Germany, the United States, and Britain. The location of infringement could be the uploader's country, which may enable citizens of countries with lax copyright laws to abuse copyright. The infringement may also occur in the server's country, potentially allowing server operators to escape liability by choosing a country with lax copyright regulation for their conduct.[69] Finally, the infringement may also occur in the downloader's country, which here would lead to application of German, United States, and British law or the law of any other country where Françoise's work was downloaded. The location of infringement may also change depending on the theory of infringement: for instance, if Bill is sued for contributory infringement, a court would have to decide if the harm centered on Bill's contribution in Britain or the direct infringement abroad. A trial court would need to sort out these issues and the process could be lengthy, costly, and complicated.

3. Summary

While the Ninth Circuit in *Subafilms* referred to the Berne Convention, the Second Circuit in *Itar-Tass* developed its choice-of-law rule based on general conflict-of-laws theory without seeking inspiration from the Berne Convention. It therefore seems that there are two main approaches: either derive a special conflict-of-laws rule for intellectual property cases from the Berne Convention or apply general conflict-of-laws rules to intellectual property cases. The following Section will examine where the European Union member states stand regarding this bifurcation.

B. The Situation in Europe

Currently, the courts of the twenty-seven member states of the European Union apply their respective national conflict-of-laws rules to copyright infringement cases.[70] This section will examine the conflict-of-laws approaches of three major players in the European Union—Great Britain, Germany, and France—as well as Belgium, a member state that stands out by recent legislative regulation of conflict of laws in intellectual property.

---

69. *See*, James Savage, *The Pirate Bay Plans to Buy Island*, LOCAL, Jan. 12, 2007, http://www.thelocal.se/6076/20070112/.

70. This situation is bound to change: The Regulation of the European Parliament and of the Council on the Law Applicable to Non-Contractual Obligations ("Rome II") will be applied after January 11, 2009 to cases that have arisen after August 20, 2007.

1. *Great Britain*

In Great Britain, case law governing copyright infringement is similar to that in the United States. From a choice-of-law perspective, one can discern two distinct periods of thought: the decisions prior to and the decisions after the introduction of the Private International Law (Miscellaneous Provisions) Act of 1995 ("PIL").[71] The first period involved a narrow application of foreign law, while the second period may provide for a broader application of foreign law.

a) Situation Before Introduction of the Private International Law (Miscellaneous Provisions) Act of 1995

Before 1995, British courts usually combined questions of jurisdiction and choice of law when evaluating the possible application of foreign law. This combination grew from tort jurisprudence, where courts rarely applied foreign law. When applied to copyright law, these principles led to a "protection vacuum" where copyright infringement with a foreign element lacked a remedy in the British courts.

Until 1995, British case law in conflict-of-laws cases relating to torts in general was decisively coined by the principle of "double actionability."[72] According to this principle,

> an act done in a foreign country is a tort and actionable as such in England, only if it is both (a) actionable as a tort according to English law, or in other words is an act which, if done, in England, would be a tort; *and* (b) actionable according to the law of the foreign country where it was done.[73]

This principle reduces choice-of-law questions to jurisdiction questions by eliminating jurisdiction where there is a difference in law.

In copyright law, the principle of "double actionability," together with the strictly territorial character of intellectual property rights, has resulted in a general tendency among British courts to apply domestic law.[74] For example, in *Def Lepp Music v. Stuart-Brown*, the Chancery Division of the

---

71. *See* Private International Law (Miscellaneous Provisions) Act, 1995, c. 42 (U.K.).

72. *See* W.R. Cornish, *Intellectual Property Infringement and Private International Law: Changing the Common Law Approach*, 4 GRUR. INT'L. 285, 286 (1996).

73. ALBERT V. DICEY ET AL., DICEY, MORRIS AND COLLINS ON THE CONFLICT-OF-LAWS 203 (14th ed. 1993) (emphasis added). The principle was originally formulated in *Phillips v. Eyre*, (1870) L.R. 6 Q.B. 1 (Exch.).

74. *See* Def Lepp Music v. Stuart-Brown, [1986] R.P.C. 273 (Ch.) (U.K.) (referring to the Copyright Act 1956).

High Court considered an alleged infringement of a British copyright by acts committed in Luxembourg and the Netherlands.[75] The court found for the defendants. In declining to award damages, the court merged the questions of jurisdiction and choice of law instead of following the fundamental conflict-of-laws two-step analysis discussed earlier.[76] Instead, the court first considered the question of applicability of the British Copyright Act and held that the territorial nature of the Act resulted in its application only within Britain. Furthermore, the judges stressed that only infringement committed in Britain would be actionable under the Act. As a consequence, the court proceeded to the issue of jurisdiction and stated, "acts done outside the United Kingdom cannot be the subject matter of an action for infringement in the English courts."[77] On the issue of choice of law, the court referred to the principle of "double actionability" as defined in Dicey and Morris and held that it would generally give "effect to the substantive law of England *(lex fori)* as opposed to the law of the place where the act is committed *(lex loci delicti)*."[78]

Application of the *lex fori* can be very tempting because of its convenience to the local court. In the Bill-Françoise hypothetical, imagine that Françoise brings her case against Bill in a Belgian court. If the Belgian court followed *Def Lepp Music*, it would apply its own law (*lex fori)*. This is the most convenient and likely most cost-effective solution since the court will not have to investigate and interpret foreign law. Furthermore, in internet copyright infringement cases, the *lex fori* approach will have the additional advantage of applying the law of a single jurisdiction as opposed to the multiple jurisdictions under the *lex loci delicti* approach.[79]

Yet, the *lex fori* solution also presents a significant disadvantage for the defendant. Generally, the plaintiff will select the forum in which to bring her infringement action. In the hypothetical, Françoise decides to consolidate the case in Belgium, her home country. Belgium, however, has no relations to the copyrights in question, except that the owner lives in Belgium. Suppose, for example, Belgium had extremely strict copyright laws compared to France and Great Britain. The fact that Françoise incidentally chose to live and sue in Belgium would result in a major disad-

75. *Id.*
76. *See* SCOLES, *supra* note 28, at 3.
77. *See Def Lepp Music*, [1986] R.P.C. 273.
78. *Id.*
79. *See supra* Part I.

vantage to Bill. Such "negative consequences" of the law of the forum state "must be taken very seriously."[80]

Generally, the parallel between *Def Lepp Music* and *Subafilms* is striking. In both cases, the courts discuss issues of choice of law under the guise of jurisdiction and territoriality. Similar to the Ninth Circuit in *Subafilms*, the *Def Lepp Music* court looks to the principle of territoriality in order to determine whether the British Copyright Act was applicable. Yet, while the court in *Subafilms* explicitly refers to the Berne Convention as a source of a choice-of-law rule, the court in *Def Lepp Music* vehemently denies any such direct effect of the Berne Convention.[81] In fact, the direct application of the Berne Convention to copyright cases is "contrary to the common law mind,"[82] due to the fact that Great Britain is a dualist system and international treaties never have a direct effect, and instead must be incorporated by legislative act.[83]

Essentially, the British approach to conflict-of-laws issues in copyright cases before the 1995 introduction of the PIL created a protection vacuum in infringement cases with a foreign element. For infringement of a United Kingdom copyright abroad, courts would resort to the territoriality principle and, thus, refuse to handle these cases as there was no extraterritorial protection of U.K. copyrights.[84] Further, U.K copyright law would not provide protection for infringements of a foreign copyright in the United Kingdom.[85] As cases involving foreign countries became more prevalent, this situation had to be addressed.

b) The Private International Law (Miscellaneous Provisions) Act 1995 and Its Implementation by the British Courts

The PIL has likely decreased the protection vacuum for foreign copyright cases in place prior to 1995. Section 10 of the PIL explicitly abolish-

---

80. Annette Kur, *Applicable Law: An Alternative Proposal for International Regulation—The Max-Planck Project on International Jurisdiction and Choice of Law*, 30 BROOK. J. INT'L L. 951, 977-78 (2005).

81. *See Def Lepp Music*, [1986] R.P.C. 273 ("I do not understand how the Berne Convention (being simply a treaty obligation) conferred any individual rights on the plaintiff under English law in the absence of any legislation incorporating the Convention into English law.").

82. Cornish, *supra* note 72, at 287 n.26.

83. *See* John H. Jackson, *Status of Treaties in Domestic Legal Systems: A Policy Analysis*, 86 AM. J. INT'L L. 310, 319 (1992).

84. *See Def Lepp Music*, [1986] R.P.C. 273; DICEY, *supra* note 73, at 1908-09.

85. *See* Copyright, Designs and Patents Act, 1988, c. 48 (U.K.); DICEY, *supra* note 73, at 1908-09.

es the "double actionability" rule.[86] Consequently, a common justification for applying the *lex fori* has been taken away.[87] Indeed, section 11(1) of the PIL provides for the general application of the *lex loci delicti* in torts.[88] Nonetheless, several sections of the PIL limit the general application of *lex loci delicti*. While not dispositive, the case law seems to hint that these exception sections do not limit the general applicability of section 11(1) in copyright infringement cases.

Three sections of the PIL potentially limit the general application of *lex loci delicti*. Section 11(2) of the PIL provides a special rule for torts involving multiple countries:

> Where elements of those events occur in different countries, the applicable law under the general rule is to be taken as being—
>
> (a) for a cause of action in respect of personal injury caused to an individual or death resulting from personal injury, the law of the country where the individual was when he sustained the injury;
>
> (b) for a cause of action in respect of damage to property, the law of the country where the property was when it was damaged; and
>
> (c) in any other case, the law of the country in which the most significant element or elements of those events occurred.[89]

Although British legislators stated that section 11(2) of the PIL did not apply to intellectual property infringement cases,[90] some scholars still advocate its use in infringements covering multiple countries.[91] In such a

---

86. *See* Private International (Miscellaneous Provisions) Act, 1995, c. 42, § 10 (U.K.). The provision reads as follows:

> The rules of the common law, in so far as they—(a) require actionability under both the law of the forum and the law of another country for the purpose of determining whether a tort or delict is actionable . . . , are hereby abolished so far as they apply to any claim in tort or delict which is not excluded from the operation of this Part by section 13 below.

Id.

87. *See supra* discussion II.B.1.a).

88. *See* Private International (Miscellaneous Provisions) Act, § 11(1) ("The general rule is that the applicable law is the law of the country in which the events constituting the tort or delict in question occur . . . .").

89. Private International (Miscellaneous Provisions) Act, § 11(2).

90. *See* PEINZE, *supra* note 1, at 350; *see also* J.J. FAWCETT & PAUL TORREMANS, INTELLECTUAL PROPERTY AND PRIVATE INTERNATIONAL LAW 621 (1998).

91. *See* FAWCETT & TORREMANS, *supra* note 90.

case, the judge would choose the law of the country in which the most significant element of infringement occurred. As is the case with the other two exception sections, English courts have yet to apply section 11(2) to cases involving intellectual property law.

The second exception is in section 14 (4) of the PIL, which provides that it shall not modify "the rules of private international law that would otherwise be so applicable."[92] The territorial nature of copyrights could trigger the application of this exception.[93] In fact, due to the territorial limitation of substantive intellectual property rules (i.e. British copyright law only protects local copyrights against local copyright infringement), they could be interpreted as mandatory rules of the forum.[94] Such mandatory rules of the forum are defined as rules of substantive law that "are regarded as so important that as a matter of construction or policy they must apply in any action before a court of the forum, even where the issues are in principle governed by a foreign law selected by choice-of-law rule."[95] As a result, British courts would always apply British copyright law. Mandatory rules of the forum most likely play a role in infringement cases, but the extent of their role is unclear.[96] The definitive decision remains with the courts.[97]

Determining the impact of the PIL on copyright infringement cases poses a challenge to British courts. A major British treatise on conflict of laws finds that "[t]he precise relationship between the protection of intellectual property rights and [the PIL] is not without difficulty."[98] Nonetheless, most courts that consider the relationship indicate that the PIL opens the door for broader application of foreign law.

In *Pearce v. Ove Arup Partnership Ltd.*,[99] the Court of Appeals provided one possible forecast for future implications of the PIL in copyright cases.[100] The case involved a British plaintiff suing several defendants, some domiciled in Britain and some in the Netherlands. The plaintiff claimed that defendants had infringed his British copyright on architectur-

92. Private International (Miscellaneous Provisions) Act, § 14(4).

93. *See* DICEY, *supra* note 73, at 1907-1908; Cornish, *supra* note 72, at 288.

94. *See* FAWCETT & TORREMANS, *supra* note 90, at 457.

95. *Id.* at 456-457 (citing Law Commission Working Paper No 87, para 4.5 (1954)).

96. *Id.* at 600.

97. Section 12 of the PIL provides an additional exception to the *lex loci delicti* in cases where application of a different law is "substantially more appropriate." Private International (Miscellaneous Provisions) Act, § 12. An application of section 12 to copyright infringement cases is, however, unlikely. *See* Cornish, *supra* note 72, at 287 n.25.

98. DICEY, *supra* note 73, at 1907.

99. *See* Pearce v. Ove Arup P'ship Ltd., [2000] Ch. 403 (A.C.) (U.K.).

100. *Id.*

al drawings when they constructed a building in Rotterdam (The Netherlands). In dicta, the court indicated that PIL section 11(1) would have led to the application of the *lex loci delicti*, thus requiring application of Dutch law.[101]

Overall, there only exist a small number of copyright cases in English courts where the question of the applicable law is explicitly examined. Most cases that involve foreign law address those issues at the jurisdiction level, and if foreign law proves decisive, are denied due to lack of jurisdiction.[102] This approach to conflict of laws in copyright law still leaves us with a protection vacuum.[103] In a more recent case, however, the Chancery Division suggested a shift, which could avoid this vacuum: in *R. Griggs Group Ltd. v. Evans*,[104] a British footwear company sued an Australian footwear company for infringement of its logo. The court suggested that, instead of refusing to deal with infringement cases involving a foreign element at all, judges should apply the conflict-of-laws rules of the 1995 Act. In finding for the British company, the court argued that

> [i]n any case the double-actionability rule was abolished by s.10 of the Private International Law (Miscellaneous Provisions) Act 1995, with the effect that, in general, it is now enough to show that the act complained of is actionable according to the law of the country where the event took place. Hence (questions of Convention, comity and forum conveniens apart) it is now sometimes possible to sue in England for infringement of a foreign intellectual property right.[105]

While this approach will not completely offer otherwise absent copyright protection to foreign copyrighted works in the United Kingdom, it will at least broaden the horizon by getting English courts to apply foreign law.[106] Such a readiness to apply foreign law has long been present among German courts.

2. Germany

Although Germany is a civil law country, German written law lacks any provision regulating the law applicable to copyright infringement cas-

---

101. *Id.*
102. *See* DICEY, *supra* note 73, at 1908.
103. See discussion *supra* Part II.A.1 on *forum non conveniens*.
104. *See* R. Griggs Group Ltd. v. Evans, [2004] EWHC 1088, [2005] Ch. 153 (U.K.).
105. *Id.*
106. *See* DICEY, *supra* note 73, at 1908-09.

es. In other words, the German Copyright Law only contains substantive rules on copyrights and related rights, but no choice-of-law rules.[107] The general choice-of-law rules in Germany do not apply to intellectual property cases.[108] This regulatory gap is due to uncontested case law by the *Bundesgerichtshof* (Federal Court of Justice). During the last revision of the choice-of-law regulations in torts, the *Deutsche Bundestag* (Federal Diet) considered a regulation of the conflict of laws in copyright infringement cases unnecessary due to the overall validity of the so-called *Schutzlandsprinzip* (law of the country for which protection is sought)[109] in the area of intellectual property law.[110] As a result of this lack of codification, it is necessary to turn to German case law, in particular copyright infringement cases decided by the *Bundesgerichtshof*.

In 1992, the *Bundesgerichtshof* indicated that the proper law for copyright infringement was the *lex protectionis*, i.e. the law of the country for which protection is sought or *Schutzlandsprinzip*.[111] According to the court, the *Schutzlandsprinzip* is founded on article 5 (2) of the Berne Convention,[112] which reads:

> The enjoyment and the exercise of these rights shall not be subject to any formality; such enjoyment and such exercise shall be independent of the existence of protection in the country of origin of the work. Consequently, apart from the provisions of this Convention, the extent of protection, as well as the means of redress afforded to the author to protect his rights, *shall be governed exclusively by the laws of the country where protection is claimed*.[113]

It has been suggested that this article 5 (2) of the Berne Convention points to the *lex protectionis*.[114] It remains unclear, however, exactly

---

107. *See* PEINZE, *supra* note 1, at 117.

108. *See, e.g.*, Bundesgerichtshof [BGH][Federal Court of Justice] Nov. 7, 2002, 152 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 317 (F.R.G.).

109. Bundesgerichtshof [BGH][Federal Court of Justice], June 17, 1992, 118 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 394 (F.R.G.).

110. *See* BTDrucks 14/343, at 10, available at http://dip.bundestag.de/parfors/parfors.htm.

111. Bundesgerichtshof [BGH][Federal Court of Justice], June 17, 1992, 118 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 394 (F.R.G.).

112. *Id.* at 397.

113. *See* Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as last revised at Paris, July 24, 1971, S. Treaty Doc. No. 99-27, 1161 U.N.T.S. 3, art. 5(2) (1986) (emphasis added).

114. *See*, e.g., Bundesgerichtshof [BGH][Federal Court of Justice] June 17, 1992, 118 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 394, 397 (F.R.G.).

which law the *lex protectionis* favors and which aspects of copyright law the *lex protectionis* governs.

With regards to the copyright infringement law applicable under the *lex protectionis*, it has been contended that this would correspond to application of the *lex loci delicti*.[115] Some even suggest that, in the end, it all boils down to application of the *lex fori*, since an infringement case usually is brought to the court of the country of protection that will apply its own law.[116] In the Bill-Françoise example, that means that if Françoise brought a suit in a German court for the infringement of her copyright in Germany, the German court would apply the *lex protectionis* which would lead to application of German law, an outcome synonymous with one stemming from application of either *lex loci delicti* or *lex fori.*

Yet, a study of German case law indicates that the *Bundesgerichtshof* does not endorse either of these two interpretations. In a 1997 case, the court applied foreign copyright law as the *lex protectionis* and, therefore, negated any efforts that would analogize the *lex protectionis* to the *lex fori*.[117] Furthermore, in 2002 the court explicitly stated that the ordinary choice-of-law rules do not apply to intellectual property cases.[118] This statement seems to imply that *lex protectionis* is distinguishable from *lex loci delicti* because *lex loci delicti* is the standard approach in general choice-of-law cases involving torts.

The extensive application of the *lex protectionis* further corroborates the difference between the *lex loci delicti* and the *lex protectionis*. According to the court, the *lex protectionis* shall govern the extent of a given in-

---

115. *See* VAN EECHOUD, *supra* note 9, at 105 ("*[L]ex protectionis* and *lex loci delicti* are often used interchangeably.").

116. *See* K. Lipstein, *Intellectual Property: Parallel Choice of Law Rules*, 64(3) CAMBRIDGE L.J. 593, 607 (2005).

117. *See* Bundesgerichtshof [BGH][Federal Court of Justice] 1997, 136 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 380, 385 (F.R.G.). The court stated:

> Das Berufungsgericht hat die Klageansprueche rechtsfehlerhaft allein nach deutschem Recht beurteilt . . . . Die Klaegerin hat ihre Ansprueche nur darauf gestuetzt, dass die Beklagte ausschliessliche Fernsehauswertungsrechte and dem Spielfilm 'Spielbankaffaere,' die ihr fuer das Gebiet von Luxemburg zustuenden, verletzt habe. Schutzland ist hier demgemaess Luxemburg. Die Klaegerin behauptet nicht, auch fuer das Gebiet der Bundesrepublik Deutschland Fernsehauswertungsrechte zu besitzen.

Id.

118. *See* Bundesgerichtshof [BGH][Federal Court of Justice] Nov. 7, 2002, 152 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 317 (F.R.G.).

tellectual property right as well as possible infringement claims.[119] Furthermore, the court also held the *lex protectionis* applicable in cases of copyright exploitation.[120] Such broad interpretation of the *lex protectionis* emphasizes its role as a special choice-of-law rule in intellectual property cases.[121] In contrast, the *lex loci delicti* is more restrictive and only applies to torts.[122] Thus, German case law implicates a major discrepancy between the *lex protectionis* and the *lex loci delicti*.

Yet despite such discrepancy, application of the *lex protectionis* to internet copyright infringement cases leads to an outcome that is as unsatisfying as the application of the *lex loci delicti*. Thus, in the Bill-Françoise hypothetical, a court endorsing the *lex protectionis* approach would have to apply the laws of each country where Françoise's work was downloaded. Such an approach would lead to the application of numerous laws and could be costly and time-consuming. Because of the ubiquitous nature of internet copyright infringement, courts would benefit greatly from the presence of a single connecting factor in order to streamline judicial procedures. Unfortunately, such rationalization has yet to be introduced into German law. Perhaps German courts can look to their French neighbors' decisions as sources of inspiration.

3. France

France has a long tradition of intellectual property protection.[123] Originally, protection in France was limited to works published in France by French authors.[124] French law gradually expanded to cover works by foreign authors and infringement in foreign territories. When considering infringement cases involving foreign elements, France's highest court, the

---

119. *See* Bundegerichtschof [BGH][Federal Court of Justice], June 17, 1992, 118 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 394, 397-98 (F.R.G.).

120. Bundesgerichtshof [BGH][Federal Court of Justice] Nov. 7, 2002, 152 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 317 (F.R.G.) ("Das Recht des Schutzlandes bestimmt, welche Handlungen als Verwertungshandlungen unter ein von ihm anerkanntes Schutzrecht fallen.").

121. Such a broad perception of the *lex protectionis*, however, is problematic in view of the wording of article 5 (2) of the Berne Convention as well as other provisions in the Convention that indicate different laws applicable to different issues of copyright. *See* SAM RICKETSON & JANE C. GINSBURG, INTERNATIONAL COPYRIGHT AND NEIGHBORING RIGHTS: THE BERNE CONVENTION AND BEYOND 1299 (2006).

122. *See* VAN EECHOUD, *supra* note 9, at 105.

123. *See* A. Lucas & H.-L. Lucas, Traite de la Propriete Intellectuelle 785 (2006); A. Chambellan, *France*, *in* Intellectual Property Laws of Europe 145, 147 (George Metaxas-Maranghidis ed., 1995).

124. *See* LUCAS & LUCAS, *supra* note 123, at 785.

Cour de cassation, has implicitly adopted a *lex loci delicti* approach.[125] Yet, it is still unclear how this *lex loci delicti* approach applies to internet copyright infringement cases. This Section will discuss several proposals directed at French courts.

The strictly territorial application of French copyright law was gradually expanded in the nineteenth century. A decree mentioned reconciling reproduction rights of foreign authors with those of their French colleagues for the first time in 1810.[126] This alignment was gradually extended to other aspects of copyright.[127] In 1852, the French government adopted a decree that extended copyright protection to works of foreign authors published abroad.[128] Interestingly, French case law in the area of copyright protection lagged behind statutory development and first consisted of narrow criminal law decisions before involving civil law issues, including choice of law.[129]

When French courts began reconciling the rights of foreign and French authors, they restricted the application of the 1852 decree to reproduction rights.[130] For example, in the *Verdi* case, the French *Cour de cassation* denied the famous Italian composer copyright protection in France.[131] The court held that in order to be protected in France, first publication of a work by a foreign author had to be in France.[132] In 1887, the *Grus* case further defined the conditions for protection of a foreign work in France.[133] According to the court in *Grus*, a foreign work had to fulfill the requirement of *double protection* in order to be protected in France: i.e. a foreign work was only protected in France if it was also protected abroad.[134] This *double protection* condition has striking similarities with the British "double actionability" rule.[135]

French courts made further progress in protecting foreign copyrights at the beginning of the twentieth century when they endorsed the principle of

---

125. *See* Cour de cassation [Cass. 1e civ.][highest court of ordinary jurisdiction], Mar. 5, 2002, Bull. civ. I, No. 75.

126. *Id.* at 736.

127. *Id.*

128. *Id.* at 786-787.

129. *Id.* at 785.

130. *Id.* at 786 and 787.

131. *See* Cour de Cassation [Cass.][highest court of ordinary jurisdiction], 1857, DP 1858, report Ferey (Fr.).

132. *Id.*

133. *See* Cour de Cassation [Cass.][highest court of ordinary jurisdiction], 1887, DP 1888, note Sarrut (Fr.).

134. *Id.*

135. *See supra* Section II.B.1.a).

reciprocity.[136] This principle made copyright protection of foreign works contingent upon the protection of French works abroad.[137] Simultaneously, the *Cour de cassation* in the *Leduc* case shifted the attention away from international criminal law aspects and, for the first time, adopted a conflict-of-laws approach dealing with civil law.[138] The courts further substantiated this new direction by holding choice-of-law rules, particularly *lex loci delicti*, applicable in copyright infringement cases.[139]

In 1959, the French *Cour de cassation* rendered a landmark decision extending French protection to a foreign copyright held by a foreign national.[140] The case involved a Russian copyright that was infringed in France by Fox-USA and Fox-Europe. The court held that the Russian authors could receive damages based on French law. It emphasized that foreigners generally enjoy the same individual rights (*droits privés*) as domestic copyright holders, unless application to foreigners is explicitly excluded.[141] As a result, the court applied French law to the infringement of the Russian copyright in France.[142]

Yet, in the *Fox* decision, the *Cour de cassation* applied French law without determining the exact choice-of-law rule applicable to copyright infringement cases. In fact, application of French law in the above case could be the result of choosing the *lex fori*, the *lex loci delicti*, or the *lex protectionis*. The French literature interpreted the approach by the *Cour de cassation* in this case as pointing towards the *lex fori*, which most scholars erroneously equated with the *lex loci delicti*.[143]

French judges, however, have leaned more towards exclusive application of the *lex loci delicti*.[144] In *Sisro*, the *Cour de cassation* upheld the application of foreign laws to infringements committed abroad and French law to infringements committed in France.[145] By upholding this judgment,

---

136. *See* LUCAS & LUCAS, *supra* note 123, at 788.

137. *Id.*

138. *Id.*

139. *Id.* at 788-789.

140. *See* Cour de Cassation [Cass. 1e civ.][highest court of ordinary jurisdiction], Dec. 22, 1959, 1960 REV. CRIT. DE DROIT INT. PRIV. 361 (Fr.).

141. *Id.* at 361.

142. *Id.* at 362.

143. *See* LUCAS & LUCAS, *supra* note 123, at 813. For an explanation of the difference between *lex fori* and *lex loci delicti*, see *supra* Part I.

144. *See* LUCAS & LUCAS, *supra* note 123, at 871.

145. *See* Cour de cassation [Cass. 1e civ.][highest court of ordinary jurisdiction], March 5, 2002, Bull. Civ. I, No. 75.

the *Cour de cassation* adopted a *lex loci delicti* approach based on article 5(2) of the Berne Convention.[146]

While the *lex loci delicti* or law of the place of the wrong appears to be the settled approach to conflict of laws in France, no court has yet resolved how this principle applies in presence of the challenge posed by internet copyright infringement cases (i.e. when there are multiple locations related to copying). Scholars in France have proposed several alternatives to the *Sisro* rule that may resolve this issue.[147]

Specifically, some have proposed to apply the law of the country where the victim lives in internet copyright infringement cases.[148] In the Bill-Françoise hypothetical, this solution would lead to application of Belgian law because Françoise resides in Belgium. Application of the law of the country where the victim lives avoids conflicts of multiple laws under *lex loci delicti*, yet has two significant drawbacks. First, such a rule would critically depend on a fact possibly unrelated to the actual case. In the hypothetical, Belgium would have no link to the case if Françoise did not live there. Second, the proposed solution unfairly favors copyright owners because they can choose their place of living.

Another suggestion has been to link choice of law to jurisdiction and follow the *lex fori* approach. According to this view, a copyright owner could choose to bring her case in any court with jurisdiction—presumably either in a country where the infringing content was posted to the Internet or in a country where the infringing content was downloaded. The court would then apply its own law.[149] In our hypothetical, under this approach, Françoise may choose whether she would like to bring her case to a French, German, United States, or British court. Françoise would probably settle on the court with the most advantageous law for her case. Such forum shopping, however, disadvantages the defendant and therefore should be avoided. In addition, the linking of jurisdiction and choice of law is contrary to basic principles of conflict of laws.[150]

Finally, it has been suggested that internet copyright infringement cases be analogized to broadcasting infringement cases.[151] The latter are regulated by the French Intellectual Property Code.[152] Under the code, broad-

---

146. *Id.*
147. *See* LUCAS & LUCAS, *supra* note 123, at 827-828.
148. *Id.* at 827.
149. *Id.* at 828.
150. *See, e.g.*, ALI, *supra* note 13, at 63.
151. LUCAS & LUCAS, *supra* note 123, at 831.
152. *See* CODE PROPRIETE INTELLECTUELLE [C. PROP. INT.] art. L122-2, *available at* http://www.legifrance.gouv.fr/html/codes_traduits/cpialtext.htm [English version].

cast infringements are governed by the law of the location where the work was transmitted to a satellite.[153] In the case of internet copyright infringement, such an approach would apply the law of the country where the infringing content was uploaded to an internet server.[154] In the hypothetical, this approach would lead to the application of British law because Bill uploaded the infringing content in Great Britain. Yet, Bill could have uploaded the work from anywhere and, under an adapted *lex loci delicti* approach, he would probably make sure that the country where he uploaded infringing content had lax copyright protection or even no protection at all (at least if he intended to infringe copyrights on a larger scale). In other words, the adapted *lex loci delicti* approach unjustifiably favors the infringer and, therefore, does not seem to be an ideal approach. Moreover, while satellites are outside all territorial jurisdiction, Bill uploaded the work to a server in a particular country. That country's laws may be better suited to deal with the infringement.

In sum, France's treatment of conflict of law issues in internet copyright infringement cases is no clearer than Germany's. A clarifying decision by the French *Cour de cassation* has yet to come down. Perhaps a look at Belgium, which historically has been a French satellite state, can shed more light on how to solve the pressing issue of determining the law applicable to copyright infringement in the internet age.

4. Belgium

Belgian law shares much with French law, but Belgian law has recently diverged from French and moved towards German law by adopting the *lex protectionis* approach to choice-of-law issues. In 2004, the Belgian legislature passed a new law codifying the *lex protectionis* approach to choice of law.[155] The new statute, however, contains several exceptions which have yet to be applied to internet copyright infringement cases. Thus, it remains to be seen how far Belgian law has strayed from its French roots.

Until recently, the conflict-of-laws approach in Belgium strongly resembled that in France. In fact, the Belgian Code Civil is based on the Code Napoleon, which underlies the current version of the French Code

153. *See id.* art L. 122-2, L. 122-2-1, L. 122-2-2.

154. *See* LUCAS & LUCAS, *supra* note 123, at 831.

155. *See* Loi portant le Code de droit international privé [CDIP] [Law establishing the Code of private international law], Moniteur Belge [Official Gazette of Belgium], July 27, 2004, p. 57344.

Civil.[156] In both countries, article 3 of the Code Civil was a key provision for private international law issues, but was limited in scope and had been vastly unchanged since 1804.[157]

In Belgium, there used to be little legislative material on choice-of-law issues to complement the rather broadly formulated Code Civil.[158] The resulting gap was sparsely filled by jurisprudence and scholarly work.[159] Overall, there had been a general tendency among Belgian jurists towards applying the *lex loci protectionis* to conflict-of-laws issues of copyright infringement. Yet, there was no written law on this subject.[160] When the lack of codification became noticeable, preparations of a legal text regulating private international law started in Belgium.[161]

The new law, the *Code de droit international privé* (Code on Private International Law), entered into force in October 2004.[162] It abolished article 3 of the Belgian Code Civil.[163] The drafters of the new Belgian Code on Private International Law drew inspiration from similar conflict-of-laws codifications in various countries in Europe.[164] Most remarkably, however, the Belgium Code of Private International Law included a provision on the law applicable to intellectual property issues.[165]

---

156. *See* Dominique d'Ambra, *La Fonction Politique du Code Civil pour la France*, *in* LE CODE CIVIL FRANÇAIS EN ALSACE, EN ALLEMAGNE ET EN BELGIQUE: REFLEXIONS SUR LA CIRCULATION DES MODELES JURIDIQUES 9, 10 (Dominique d'Ambra et al. eds., 2006).

157. CODE CIVIL [C. CIV.] art. 3 (Belg.). states: "Les lois de police et de sûreté obligent tous ceux qui habitant le territoire. Les immeubles, meme ceux possédés par des étrangers, sont régis par la loi belge [française]. Les lois concernant l'état et la capacité des personnes régissent les Belges [Français], même résidant en pays étranger." Despite such textual similarity, French and Belgian courts interpreted Code civil provisions differently. *See* d'Ambra, *supra* note 156, at 18). In Belgium, this provision has been recently abolished.

158. *See* CE avis no. 2-1225/1, February 12, 2001 (Belg.), *available at* www.ipr.be. In fact, the provisions of the Code Napoléon were deliberately formulated in a broad way in order to allow for a flexible contemporary interpretation by the courts. *See* D'Ambra, *supra* note 156, at 18.

159. *See* Ferdinand Visscher & Benoit Michaux, Precis du Droit d'Auteur et des Droits Voisins 631 (2000).

160. *Id.*

161. *See* François Rigaux & Marc Fallon, Droit International Privé 71 (2005).

162. *See* Loi portant le Code de droit international privé [CDIP] [law establishing the Code of private international law] of July 16, 2004, Moniteur Belge [Official Gazette of Belgium], July 27, 2004, p. 57344.

163. *Id.* p. 57373, art. 139.

164. *See* RIGAUX, *supra* note 161, at 71-72.

165. *See* CDIP, Moniteur Belge [Official Gazette of Belgium], July 27, 2004, p. 57363, art. 93.

Article 93 of the Code on Private International Law sets the *lex protectionis* as the general rule for intellectual property rights infringement cases, but also lays out several exceptions.[166] For example, article 93 provides an exception for cases where parties have stipulated by contract to a different law.[167] Another possible exception of the rule as provided in article 93 is the *ordre public* exception.[168] This exception enables a judge to avoid application of a foreign law if the law would run contrary to fundamental rules of Belgian intellectual property law.[169] It remains to be seen how likely Belgian courts will be to resort to the *ordre public* exception in order to force application of Belgian intellectual property law.

Another possible exception, article 19, provides for application of a law more closely connected to the case at issue than the law applicable according to the general rules of the Code on Private International Law.[170]

---

166. *Id.* The law provides :

> Les droits de propriété intellectuelle sont régis par le droit de l'Etat pour le territoire duquel la protection de la propriété est demandée. Toutefois, la détermination du titulaire originaire d'un droit de propriété industrielle est régie par le droit de l'Etat avec lequel l'activité intellectuelle présente les liens les plus etroits. Lorsque l'activité a lieu dans le cadre de relations contractuelles, il est présumé, sauf preuve contraire, que cet Etat est celui dont le droit est applicable à ces relations.

Id.

167. *Id.*

168. *Id.* p. 57347, art. 21. The law reads as follows:

> L'application d'une disposition du droit étranger désigné par la présente loi est écartée dans la mesure où elle produirait un effet manifestement incompatible avec l'ordre public. Cette incompatibilité s'apprécie en tenant compte, notamment, de l'intensité du rattachement de la situation avec l'ordre juridique belge et de la gravité de l'effet que produirait l'application de ce droit étranger. Lorsqu'une disposition du droit étranger n'est pas appliquée en raison de cette incompatibilité, une autre disposition pertinente de ce droit ou, au besoin, du droit belge, est appliquée.

Id.

169. *See Proposition de loi portant le Code de droit international privé*, [Draft law on the Code of private international law], 3-27/1 SE (2003) (submitted by Ledouc et al.), p. 120 (Belg.).

170. Article 19 of the CDIP, reads as follows:

> Le droit désigné par la présente loi n'est exceptionnellement pas applicable lorsqu'il apparaît manifestement qu'en raison de l'ensemble des circonstances, la situation n'a qu'un lien très faible avec l'Etat dont le droit est désigné, alors qu'elle présente des liens très étroits avec un autre Etat. Dans ce cas, il est fait application du droit de cet autre Etat. Lors de l'application de l'alinéa 1er, il est tenu compte notamment : - du besoin de prévisibilité du droit applicable, et - de la circonstance que la relation en cause a été établie régulièrement selon les règles de droit in-

Such a rule may avoid both the application of the *lex protectionis* and its detrimental effect on the holders of copyrights in cases of internet copyright infringement. It does not, however, determine precisely which rule should apply instead. In the Bill-Françoise-hypothetical, the relevant question is which law is more closely connected to the case than the *lex protectionis*? Belgian law because Françoise lives in Belgium and brings her case in front of a Belgian court? British law because Bill uploaded the infringing content while he was in Great Britain? French law because the work was (first) published in France? The exact implications of article 19 on choice of law in internet copyright infringement cases needs to be further examined by the Belgian courts.

5. Summary

This section of the Article has demonstrated, on the basis of several European member states, that European approaches are divided between the *lex protectionis* and the *lex loci delicti*. This is especially true for cases of internet copyright infringement. While some countries have not yet expressed the approach they favor (e.g., Germany, England and France), Belgium has been more explicit. Thus, Belgium will possibly end up deciding internet copyright infringement cases that involve several laws on a case-by-case basis and apply the law that is most closely connected to the case. Given the staggering number of copyright infringements on the Internet transcending territorial palladia, the current prevailing ambiguity in approaches is no longer viable.

III. TOWARDS HARMONIZATION OF CONFLICT OF LAWS IN INTERNATIONAL INTELLECTUAL PROPERTY LAW

The need for convergence on the law applicable to copyright infringement issues has precipitated two movements toward harmonization of this increasingly important area of law. On the one hand, in the European Union, such efforts have been part of a larger endeavor to create a uniform choice-of-law regime for torts. The *Regulation of the European Parliament and of the Council on the law applicable to non-contractual obligations* (Rome II) also provides for cases of intellectual property infringement.[171] On the other hand, the American Law Institute has drafted *Intellectual Property: Principles Governing Jurisdiction, Choice of Law, and*

---

ternational privé des Etats avec lesquels cette relation présentait des liens au moment de son établissement.

CDIP, Moniteur Belge [Official Gazette of Belgium], July 27, 2004, p. 57347, art. 93.

171. *See* Rome II, *supra* note 12, art. 8.

Judgments in Transnational Disputes (ALI Principles), a comprehensive regulation of conflict-of-laws issues specifically focused on intellectual property rights.[172]

It is remarkable that 2007 has seen the adoption of two distinct international instruments, the ALI Principles and Rome II that provide, among other things, choice-of-law rules for cases of intellectual property infringement. Rome II is the long expected harmonization of European choice-of-law rules in the area of non-contractual relationships in general. It includes, however, one specific provision relating to choice-of-law issues in intellectual property infringement. Rome II has the form of a European Union regulation. As a result, it is "binding in its entirety and directly applicable in the Member States in accordance with the Treaty establishing the European Community."[173] The ALI Principles, on the other hand, are recommendations regarding the conflict-of-laws aspects of intellectual property rights cases in particular. Though they are not binding, the ALI Principles serve as a guide that could equip courts and legislators to better deal with complex choice-of-law issues pertaining to intellectual property.[174] This Part will analyze, evaluate, and compare both harmonization efforts.

A. Rome II—Harmonization of Choice-of-Law Rules Relating to Torts

In Europe, the drafting of Rome II has taken place over a period of several years. In 2002, the European Commission first presented a preliminary draft of a regulation on choice of law in torts and thereby opened a period of discussion.[175] During this discussion period, Rome II's extension to intellectual property rights cases was particularly contested.[176] Oppo-

---

172. *See* ALI, *supra* note 13.

173. Rome II, *supra* note 171, at 48 (closing sentence); Consolidated Version of the Treaty Establishing the European Community, art. 249(2), Dec. 24, 2002, 2002 O.J. (C 325) 33, 65 [hereinafter EC Treaty].

174. *See* François Dessemontet, *A European Point of View on the ALI Principles—Intellectual Property: Principles Governing Jurisdiction, Choice of Law, and Judgments in Transnational Disputes*, 30 BROOK. J. INT'L L. 849, 855 (2005).

175. *See* Consultation on a Preliminary Draft Proposal for a Council Regulation on the Law Applicable to Non-Contractual Obligations (May 2002), *available at* http://ec.eur-opa.eu/justice_home/news/-consulting_public/rome_ii/printer/news_hearing_rome2_en.htm (last visited May 21, 2008) [hereinafter Rome II Consultation].

176. *See* Summary and contributions of the consultation "Rome II": Follow up on 'Consultation on a Preliminary Draft Proposal for a Council Regulation on the Law Applicable to Non-Contractual Obligations ("Rome II")' (Oct. 2002), available at http://ec.europa.eu/justice_home/news/consulting_public/rome_ii/news_summary_rome2_en.htm (last visited May 21, 2008) [hereinafter Rome II Contributions].

nents argued that intellectual property law was dominated by the principle of territoriality that warrants special rules addressed in a separate instrument on intellectual property law.[177] As a result, major revisions have been made and a separate provision for intellectual property infringements has been added to Rome II.[178] Section 1 examines the *travaux préparatoires* (the preparatory work and the official record of negotiation) of Rome II, in particular the regime relating to intellectual property rights. An in-depth analysis of the current version of Rome II is undertaken in Section 2, followed by an evaluation of the current text and its implications for future copyright infringement cases.

1. Travaux Préparatoires

In May 2002, the European Commission first proposed the text for a planned regulation applicable to non-contractual obligations.[179] The regulation was part of an effort to facilitate the free movement of persons[180] by enhancing judicial cooperation in civil matters.[181] The proposed preliminary draft of Rome II was a sequel to previous European harmonization efforts in the area of conflict of laws, in particular the "Brussels I" Regulation, which harmonized the areas of jurisdiction, recognition and execu-

---

177. *See, e.g.*, Comment of European Federation of Pharmaceutical Industries and Associations on the 'Commission Draft Proposal for a Council Regulation on the Law Applicable to Non-Contractual Obligations ("Rome II") (Sept. 27, 2002) (for a separate regulation of choice of law in intellectual property cases), *available at* http://ec.europa.eu/justice_home/news/consulting_public/rome_ii/euro_feder_pharma_en .pdf. *But see* Comment of Institut für ausländisches und internationales Privat-und Wirtschafsrecht der Universität Heidelberg on the 'Commission Draft Proposal for a Council Regulation on the Law Applicable to Non-Contractual Obligations ("Rome II")' (Sept. 14, 2002), *available at* http://ec.europa.eu/justice_home/news/consulting_public/ rome_ii/universitat_ heidelberg_de.pdf.

178. *See* Commission Proposal for a Regulation of the European Parliament and the Council on the Law Applicable to Non-Contractual Obligations ("Rome II"), COM (2003) 427 final, art. 8 (July 22, 2003), *available at* http://eur-lex.europa.eu/LexUriServ/ LexUriServ.do?uri=COM:2003:0427:FIN:EN:PDF [hereinafter Rome II Proposal].

179. *See* Rome II Consultation, *supra* note 175.

180. *See* Consolidated Version of the Treaty on European Union, art. 2, Dec. 24, 2002, 2002 O.J. (C 325) 5 ("The Union shall set itself the following objectives: . . . to maintain and develop the Union as an area of freedom, security and justice, in which the free movement of persons is assured in conjunction with appropriate measures . . . .").

181. *See* EC Treaty, *supra* note 173, art. 61(c) ("In order to establish progressively an area of freedom, security and justice, the Council shall adopt . . . measures in the field of judicial cooperation in civil matters as provided for in Article 65(b) . . . ."); *id.* art. 65 ("Measures in the field of judicial cooperation in civil matters having cross-border implications . . . shall include: . . . promoting the compatibility of the rules applicable in the Member States concerning the conflict of laws and of jurisdiction . . . .").

tion of judgments in contracts, torts and trade law,[182] and the "Rome I" Convention, which unified choice-of-law rules in contracts law.[183]

In the preliminary draft of Rome II, intellectual property infringement was neither excluded nor specially addressed.[184] Consequently, choice-of-law questions on intellectual property infringement cases fell into the scope of application of the general choice-of-law provision, which read as follows:

> The law applicable to a non-contractual obligation arising out of a tort or delict shall be the law of the country in which the loss is sustained, irrespective of the country or countries in which the harmful event occurred and irrespective of the country in which the indirect consequences of the harmful event are sustained, subject to paragraph 2.
>
> Where the author of the tort or delict and the injured party have their habitual residence in the same country when the tort or delict is committed, the applicable law shall be the law of that country.
>
> However, if it appears from the circumstances as a whole that there is a substantially closer connection with another country and there is no significant connection between the non-contractual obligation and the country whose law would be the applicable law under paragraphs 1 and 2, the law of that other country shall be applicable.
>
> A substantially closer connection with another country may be based in particular on a pre-existing relationship between the parties, such as a contract that is linked to the tort or delict in question.[185]

The following paragraphs will provide a section-by-section discussion of article 3 and its application to internet copyright infringement cases.

---

182. Council Regulation 44/2001, 2001 O.J. (L 12) 1 (EC).

183. Convention on the Law Applicable to Contractual Obligations 80/934/EEC, *opened for signature* June 19, 1980, 1980 O.J. (L 266) 1.

184. The draft offered the following general provision: "The rules of this Regulation shall apply to non-contractual obligations in any situation involving a choice between the laws of different countries." Rome II Consultation, *supra* note 175, art. 1(1). It then offered a list of excluded obligations, none of which included intellectual property. Rome II Consultation, *supra* note 175, art. 1(2).

185. Rome II Consultation, *supra* note 175, art. 3.

a) Article 3 Section 1: Lex loci delicti

According to article 3(1) of Rome II, the *lex loci delicti* governs choice-of-law issues arising in intellectual property infringement cases.[186] This applicability of the *lex loci delicti* to intellectual property infringement issues has been subject to major criticism, in particular from intellectual property and conflict-of-laws scholars. One contention against the *lex loci delicti* rule was that it had been common state practice to apply the *lex protectionis* to cases of intellectual property infringement.[187] Such argumentation is flawed because the suggested international endorsement of the *lex loci protectionis* does not exist.[188] This argument alone would therefore not suffice to prevent a *lex loci delicti* approach.

A more substantial argument against application of the *lex loci delicti* to intellectual property infringement issues is that the European Commission seemed to interpret the *lex loci delicti* as leading to the application of the law of the place where the direct damage occurred.[189] With this approach, the Commission apparently intended to accommodate tort victims, as the state of direct damage generally coincided with their country of residence.[190] Yet this is not at all advantageous for copyright infringement victims. In fact, due to the territoriality of copyrights, direct damage would be in the country where the infringing act occurred.[191] Moreover, as explained above, application of the *lex loci delicti* to internet copyright infringement issues would result in the application of multiple laws.[192] Given these substantive disadvantages, the *lex loci delicti* approach was inadvisable and therefore abandoned in subsequent drafts of Rome II.[193]

b) Article 3 Section 2: Law of the habitual residence

Another possible problem with respect to intellectual property rights was that the preliminary draft version of article 3(2) provided for the law of the habitual residence if both parties were residents of the same place.[194] This rule is contradictory to the principle of territoriality which governs intellectual property law. Specifically, the underlying rationale of article 3(2)—the balancing of the interests involved and affected—

186. *See id.* art. 3(1).
187. *See, e.g.*, Rome II Contributions, *supra* note 176, at 24-25.
188. *See supra* Part II.
189. *See* Rome II Proposal, *supra* note 178, at 11.
190. *Id.*
191. *See* Drexl, *supra* note 15, at 154.
192. *See supra* Section II.A.2.
193. *See, e.g.*, Rome II Proposal, *supra* note 178, art. 8(1).
194. *See* Rome II Consultation, *supra* note 175, art. 3(2).

conflicts with the strictly territorial character of intellectual property rights.[195] Yet, such a deviation from common principles of intellectual property law may be justified because it simplifies lawsuits, particularly in cases of ubiquitous infringement of copyrights on the Internet.

c) Article 3 Section 3: Law of the country of the substantially closer connection

An alternative exception to applying the general *lex loci delicti* rule to copyright infringement cases could have been provided by article 3(3) of the preliminary draft. This section provided that if a case had a substantially closer connection with a country other than the one selected by *lex loci delicti*, the law of the other country would be applicable.[196] One could argue that the law of the country for which protection is sought would present a "substantially closer connection" in copyright infringement cases.[197] Yet, to avoid uncertainty and differing interpretation by the courts of the European member states, the European Commission cautioned in its revised 2003 proposal that the application of section 3 should be only under exceptional circumstances.[198] Consequently, the repeated application of the *lex loci protectionis* in copyright infringement cases was not an option.

Overall, article 3 of Rome II seems inadequate when it comes to choice-of-law issues in intellectual property cases. Accordingly, critics demanded an article specific to intellectual property infringements that would consider the particularities of intellectual property. The European Commission reacted to these criticisms by adding an article pertaining to infringement of intellectual property rights: Article 8.[199]

2. *Analysis of the Current Version of Rome II*

After the initial consultation phase, the European Commission revised its original proposal and incorporated many propositions made by experts and other stakeholders. As for intellectual property infringement, the Commission largely adopted one version of a specific intellectual property article that was proposed by a group of experts.[200] Article 8(1) of the revised Rome II Regulation provides the *lex loci protectionis* as the law applicable to intellectual property rights infringement cases in general. Ar-

195. *See* Drexl, *supra* note 15, at 155.
196. *See* Rome II Consultation, *supra* note 175, art. 3(3).
197. *See* Drexl, *supra* note 15, at 155.
198. *See* Rome II Proposal, *supra* note 178, at 12.
199. *See id.* art. 8; Rome II, *supra* note 171, art. 8.
200. *See* Rome II Contributions, *supra* note 176.

ticle 8(2) sets forth a specific rule for community intellectual property rights and article 8(3) explicitly excludes party autonomy for cases of intellectual property rights infringement. Beginning with the introduction, article 8(1) reads as follows:

> The law applicable to a non-contractual obligation arising from an infringement of an intellectual property right shall be the law of the country *for which* protection is claimed.[201]

The text of article 8(1) of Rome II has slightly different wording than article 5(2) of the Berne Convention which provides that:

> [T]he extent of protection, as well as the means of redress afforded to the author to protect his rights, shall be governed exclusively by the laws of the country *where* protection is claimed.[202]

Thus, in Rome II the *where* of the Berne Convention has been replaced by *for which*. The reason for this difference is much more than just a matter of style. Instead, with this new language the drafters of Rome II intended to avoid the common confusion that resulted from the ambiguous wording of article 5(2) of the Berne Convention as a choice-of-law rule. In fact, a literal interpretation of article 5(2) would suggest the *lex fori*, i.e. the law of the country where the plaintiff has filed his complaint.[203] Yet, the country of the forum may not have a connection with the copyright at issue: a court may have been chosen merely because the defendant has assets in the forum state, when the copyright infringement occurred in another state. There is no reason for application of the law of the forum state in such a case.[204] Consequently, the literary interpretation of article 5(2) of the Berne Convention was largely avoided and the clause was interpreted as pointing towards application of the *lex protectionis*. Some European Member States therefore will have to reassess their approach to choice-of-law issues in intellectual property infringement cases. Yet, as mentioned above, the *lex protectionis* proves to be problematic in internet copyright infringement cases.[205]

---

201. Rome II, *supra* note 171, art. 8(1) (emphasis added).
202. Berne Convention, *supra* note 17, art. 5(2) (emphasis added).
203. *See* VAN EECHOUD, *supra* note 9, at 103.
204. *See supra* Section II.B.1.a); *see also* VAN EECHOUD, *supra* note 9, at 103-05.
205. *See supra* Section II.B.2.

3. *Evaluation of Article 8(1) of the Rome II Regulation*

That article 8 of Rome II does not provide a special rule for internet copyright infringement cases is deplorable. The omission of a choice-of-law regulation for internet copyright infringement conflicts with Rome II's general goals, i.e. a European harmonization of choice-of-law rules for torts. As noted above, Member States differ substantially on the issue of what law should be applied to internet copyright infringement cases.[206] Some endorse the idea of making internet copyright infringement cases more economically efficient by choosing the application of one single law to those cases—either the *lex fori* or the law of the country where the infringing act was committed (perhaps best called a broad conception of *lex protectionis)*.[207] Others, however, continue to apply the law of the country where the infringing act had its effect (perhaps best called a narrow conception of *lex protectionis*) even if that results in the application of multiple laws by the court seized.[208]

Under the current article 8 of Rome II, each approach is plausible. Proponents of the *lex fori* probably have the most difficulty explaining their conformity with article 8(1) of Rome II, but might argue that Rome II creates a vacuum to be filled by the most suitable law, i.e. the *lex fori*. In contrast, proponents of the broad and narrow conception of the *lex protectionis* can ground their contentions in article 8(1) of Rome II and its endorsement of the *lex loci protectionis* approach. Either approach results in a sufficient connection to the tort committed. Yet one resounding argument for application of the law of the country where the infringing content was uploaded is that it only points to the law of one country and is therefore more economically efficient for copyright infringement plaintiffs.

All of the above-described approaches are possible under the current version of Rome II. And considering the legal situation in the countries described above, such indeterminacy could lead to diverging jurisprudence on this matter in the various Member States, an outcome that the drafters of Rome II specifically aimed to avoid.

To avoid indeterminacy, a special provision on infringement cases that involve multiple laws would have been prudent. The final Sections of this Article will suggest a cascading approach to internet copyright infringement cases. A possible decision cascade could look like this: first to the law of the common habitual residence; if the parties do not have a common habitual residence, then second to the law of the country where the

206. *See supra* Section II.B.
207. *See supra* Section II.B.3 (France).
208. *See supra* Section II.B.4 (Belgium).

parties had a prior relationship; and if the relationship is not closely connected to the case at hand, then finally to the laws of the States towards which the parties primarily directed their activities.[209]

4. Comparing Rome II to Itar-Tass

A comparison of Rome II with *Itar-Tass*, the prevailing approach to the issue of choice of law in intellectual property infringement in the United States,[210] demonstrates the Berne Convention's failure to clearly resolve conflicts of law. In *Itar-Tass*, the Second Circuit negated any choice-of-law implications of Berne.[211] The Second Circuit emphasized that it was not bound to any rule that may flow out of the Berne Convention. In fact, the court resorted to general choice-of-law rules, i.e. the laws applicable to torts in general instead of rules specially developed for intellectual property. The court thus achieved application of the *lex loci delicti*.[212] In Rome II, however, the European Commission asserted that the Berne Convention was based on the principle of the *lex loci protectionis*.[213] Such discrepancy in approaches between the United States and Europe is unfortunate. Contrary interpretations of the Berne Convention counteract the very essence of a treaty meant to harmonize a multitude of differing copyright regimes. Therefore alignment of the two approaches should be the ultimate goal.

B. ALI Principles—Harmonization of Conflict-of-laws Rules in Intellectual Property Law on a Global Level?

The adoption of the ALI Principles could result in such alignment of the United States and European approaches.[214] Efforts to harmonize conflict-of-laws rules governing intellectual property rights on an international level advanced in 2001 when the American Law Institute (ALI) became interested in the project.[215] Since then, legal scholars and other experts from the United States and abroad have presented several draft proposals of the ALI Principles.[216] In May 2007, ALI members approved the pro-

---

209. For further information on the cascading approach to internet copyright infringement cases, *see infra* Section II.B.3.

210. *See supra* Section II.A.2.

211. *See supra* Section II.A.2.

212. *See* Itar-Tass Russian News Agency v. Russian Kurier, Inc., 153 F.3d 82, 90-91 (2d Cir. 1998).

213. *See* Rome II Proposal, *supra* note 178, at 20 (cmt. to art. 8).

214. *See* ALI, *supra* note 13.

215. *See* Dessemontet, *supra* note 174, at 850.

216. *See* Rochelle Dreyfuss, *The ALI Principles on Transnational Intellectual Property Disputes: Why Invite Conflicts?*, 30 BROOK. J. INT'L L. 819, 820 (2005).

posed final draft of the ALI Principles. The official text was recently published.[217] The following Sections briefly explain the nature of the ALI Principles, followed by a discussion of the provisions of the ALI Principles that relate to copyright infringement, and will conclude with an evaluation of the ALI Principles.

1. Nature of the ALI Principles

In contrast to Rome II, the ALI Principles lack authoritative force[218] and do not constitute a Restatement.[219] They merely propose a set of rules that courts, scholars, and lawyers can resort to in questions on conflict-of-laws issues in intellectual property law. They are aimed at supplementing, rather than changing, national law,[220] and are guided by a vision of cooperation among courts.[221] Also, unlike Rome II, the ALI Principles do not regulate conflict of laws in general, but focus on conflict-of-laws issues in intellectual property law.[222]

Yet, close cooperation during the drafting process between intellectual property scholars and conflict-of-laws scholars becomes clear in the dogmatic precision in which the principles are formulated. For example, the ALI Principles clearly state that they refer to the substantive law of a state and not to its choice-of-law rules.[223] Such clarification is welcome, because choice-of-law rules may possess different functional attributes: they may either refer to the substantive or choice-of-law rules of a given country.[224] The ALI Principles' explicit reference to substantive law therefore advances legal certainty in conflict-of-laws cases. It also prevents *renvoi*, i.e. the referring of an issue back and forth between different choice-of-law rules.[225] The structure of the ALI Principles propagates a clear division between jurisdiction and choice-of-law questions by consecrating each a separate part.[226] In section 103 of the ALI Principles, the drafters

---

217. *See* The American Law Institute, http://www.ali.org/ (last visited Aug. 4, 2008).

218. *See* Dessemontet, *supra* note 174, at 855.

219. *Id.*

220. *Id.* at 855-856.

221. *See* ALI, *supra* note 13 at 7.

222. *Id.* at xx.

223. *Id.* at 196. Rome II contains in Article 24 a similar provision that is less remarkable as the whole regulation focused on general choice-of-law theory. Prior case law in the area of international intellectual property, largely connived in the possibility of *renvoi*—if it touched the issue of choice of law at all.

224. *See* RUSSELL J. WEINTRAUB, COMMENTARY ON THE CONFLICT OF LAWS 37 (2006) (illustrative case raising that issue).

225. *See* SCOLES ET AL., *supra* note 28, at 134.

226. *See* ALI, *supra* note 13. Part II regulates jurisdiction issues and Part III provides rules for questions of applicable law. *Id.*

explicitly emphasize the distinction between questions of jurisdiction and choice-of-law issues:

> (1) Competence to adjudicate does not imply application of the forum State's substantive law.
>
> (2) A court should not dismiss or suspend proceedings merely because the dispute raises questions of foreign law.

The drafters of section 103, therefore, clearly reject decisions similar to *Subafilms* or *Def Lepp Music*.[227] Instead, questions of jurisdiction and choice of law should be considered separately without being interrelated. Apart from this remarkable dogmatic precision, the ALI Principles also entail improvements for choice-of-law rules in copyright infringement cases, which will be examined in the following section.

2. *Analysis of the Final Draft of the ALI Principles*

The ALI Principles build upon the territoriality principle even for complex multinational cases.[228] Section 301 provides:

> 1) Except as provided in §§ 302 and 321-323, the law applicable to determine the existence, validity, duration, attributes, and infringement of intellectual property rights and the remedies for their infringement is:
>
> . . .
>
> (b) for other intellectual property rights, the law of each State *for which protection is sought* . . . .[229]

According to this provision, the general rule for copyright infringement cases is the *lex protectionis*. Notably, the provision also uses the clarified expression "for which" instead of "where."[230] In this regard, the drafters of the ALI Principles explicitly refer to Rome II.[231] However, in contrast to the European Commission, the Reporters' notes specify that the Berne Convention (and other international instruments) would not imply a

---

227. *See supra* Sections II.A.1, II.B.1.
228. *See* ALI, *supra* note 13, at 3, 193-194.
229. *Id.* § 301 (emphasis added).
230. This language is the same as that of article 8(1) of Rome II, *supra* note 171. *See supra* Section III.A.2.
231. *See* ALI, *supra* note 13, at xx.

choice-of-law rule.[232] Further, the spectrum of the *lex protectionis* of the ALI Principles is much broader than that of Rome II. While the latter is only concerned with infringements, section 301 applies to issues of existence, validity, duration, attributes, and infringement.[233] Section 301 of the ALI Principles thus merely provides for the general rule. The exceptions to this rule are contained in section 302 (party autonomy exception), section 321 (ubiquitous infringement exception), section 322 (*ordre public* exception), and section 323 (mandatory rules exception). The next two Sections focus on two exceptions that are of particular interest in the context of copyright infringement on the Internet: party autonomy and ubiquitous infringement.

a) Party Autonomy Exception (section 302)

The ALI Principles allow the parties involved to determine the law applicable to their case:

> (1) Subject to the other provisions of this Section, the parties may agree at any time, including after a dispute arises, to designate a law that will govern all or part of their dispute. . . .
>
> (3) Any choice-of-law agreement under subsection (1) may not adversely affect the rights of third parties. . . .[234]

According to this rule, parties may choose the law even after their dispute has arisen, provided they do not harm third parties with their choice. To leave to the parties the determination of the law applicable to copyright infringements may prove useful for cases of multi-state infringement (e.g., through the Internet), as the parties could conceivably agree on one single law to be applicable on their case. Yet, in cases of copyright infringement, one choice of law will often benefit one party much more than the other.

Thus, in the hypothetical, Françoise and Bill might simply agree on the application of German law, for instance, because the first infringing downloads occurred in Germany. But what if Germany's copyright law was particularly favorable to copyright owners? *Ex post* party autonomy in copyright infringement cases would prove impracticable because parties simply may not agree on a choice of law as a result of their diverging interests. In the hypothetical, Bill (the copyright infringer) might insist on application

---

232. *Id.* at 208.

233. *Compare id.* § 301 (general rule for law applicable to existence, validity, duration, attributes, and infringement of intellectual property rights), *with* Rome II, *supra* note 171, art. 8(1) (solely determining law for intellectual property infringement).

234. ALI, *supra* note 13, § 302.

of the law of the country with the most lenient copyright law, whereas Françoise (the copyright owner) might prefer application of the law of a country with very strong copyright protection. Given the diverging interests of parties on the choice of the law applicable to their case, party autonomy will likely not prove helpful in internet copyright infringement cases.

b) Ubiquitous Infringement Exception (section 321)

None of the above-discussed legal regimes presented a fully satisfactory solution for cases of copyright infringement on the Internet. As such, it is remarkable that the ALI Principles are the first to provide an explicit rule for cases of ubiquitous infringement. Section 321 reads as follows:

> (1) When the alleged infringing activity is ubiquitous and the laws of multiple States are pleaded, the court may choose to apply to the issues of existence, validity, duration, attributes, and infringement of intellectual property rights and remedies for their infringement, the law or laws of the State or States with close connections to the dispute, as evidenced, for example, by:
>
> (a) where the parties reside;
>
> (b) where the parties' relationship, if any, centered;
>
> (c) the extent of the activities and the investment of the parties; and
>
> (d) the principal markets toward which the parties directed their activities.[235]

According to this provision, the judge chooses the law applicable to a case involving multiple laws. In making such choice, the court must determine the law with the closest connection to the case. Section 321 enumerates examples to aid in determining the law with the closest connection and thus the law applicable.

When examining this list, it becomes clear that the ALI Principles do not—at least primarily and explicitly—endorse the law of the country where the infringing work was uploaded (country of origin), although such an approach is taken by some countries and advocated by some scho-

235. *Id.* § 321.

lars.[236] The Reporters' Notes of the ALI Principles discard this type of approach for two legitimate reasons. First, the country of origin might be hard to determine due to the technical complexity of digital information networks.[237] Second, the law of the country where the infringing work was uploaded could unduly benefit a savvy copyright infringer. An infringer could choose a country where copyright protection is relatively weak and upload an infringing work from that country.[238]

Consequently, the ALI Principles endorsed their unique approach. In the case of both parties being citizens and residing the same country, the court would be allowed to apply the law of their home country and not of the countries where the infringement may have had its effect. It has been argued that application of the law of the common habitual residence of the parties constituted a major deviation from the general territoriality principle. It may, however, be justified for practicability reasons.[239] Generally, application of section 321 will probably be less an exception to the territoriality principle.[240] In fact, it will lead to the application of either the law of the country where the infringement inflicts most harm or the law of the country where the infringing conduct originates (if at all determinable).[241]

If, however, both of these laws differ considerably, the choice may be difficult to make and may be perceived as arbitrary.[242] Furthermore, if the best law for the case is not easily determinable, the court will likely lean towards applying its own law (*lex fori)*.[243] Application of the *lex fori*, however, would favor the copyright holder as he will most likely have chosen the court.[244]

Overall, the current version of section 321 of the ALI Principles may not prove sufficient in some cases for copyright infringement on the Internet. The introductory note to Part III of the ALI Principles on the Applicable Law offers a response to this, stating that "the Principles endeavor to set a broad and open-ended framework, rather than, perhaps prematurely, devising a full repertory of specific rules."[245] Thus, the provisions aimed at ubiquitous infringement are flexible enough to evolve with time and

236. *See, e.g.*, William Patry, *Choice of Law and International Copyright*, 48 AM. J. COMP. L. 383, 457 et seq. (2000); *see supra* Section II.B.3 (discussion on France).
237. *See* ALI, *supra* note 13, at 248.
238. *Id.*
239. *See supra* Section III.A.1.b).
240. *See* Dreyfuss, *supra* note 216, at 843-844.
241. *See* Kur, *supra* note 80, at 977.
242. *Id.* at 977-978.
243. *Id.*
244. *See supra* Section II.B.1.a).
245. ALI, *supra* note 13, at 195.

new technological developments. Yet, such flexibility is not necessarily helpful when it comes to streamlining choice-of-law rules for cases of internet copyright infringement.[246]

c) Comparing Rome II and the ALI Principles

With regard to the proper choice-of-law regime for cases of copyright infringement, the basic provisions of both Rome II and the ALI Principles are consistent. Article 8(1) of Rome II and section 302(1)(b) of the ALI Principles provide for application of the *lex protectionis*, in its unambiguous form.[247] Yet, while Rome II provides only a special rule for community rights and explicitly excludes party autonomy, the ALI Principles provide for several exceptions to the general *lex protectionis* rule, among other things, party autonomy and ubiquitous infringement.[248] While it is debatable whether a party autonomy provision should have been included in Rome II,[249] the complete omittance of a rule on ubiquitous infringement is deplorable. It remains to be seen how the ALI Principles will serve their purpose, particularly when it comes to multi-state infringement involving numerous national laws.[250]

3. Will Courts Adopt the ALI Principles?

The previous Section analyzed and discussed major provisions of the ALI Principles pertaining to the law applicable to intellectual property cases. This Section will evaluate the ALI Principles, with an emphasis on a critical question: whether they are likely to be adopted by courts. To determine the answer, this Article will consider the parallels and differences between the ALI Principles and the intellectual property provision of Rome II. This will lead to a discussion of the question of whether the ALI Principles will serve their purpose.

As discussed above, the ALI Principles do not provide a perfect solution for cases of copyright infringement on the Internet because the drafters of the ALI Principles eventually settled on a stricter concept of territoriality than originally envisioned.[251] This attenuation occurred mainly for two reasons. First, adoption of a more traditional territoriality approach preserves state sovereignty interests. Second, and a more or less direct consequence of the first point, the endorsement by the ALI Principles of

246. For more on the question whether courts will adopt the ALI Principles, see *infra* Section III.B.3.

247. *See supra* Sections III.A.3, III.B.2.

248. *See* Rome II, *supra* note 171, art. 8(2); ALI, *supra* note 13, §§ 302, 321.

249. *See supra* Section III.B.2.a).

250. *See supra* Section III.B.2.b).

251. *See* Dreyfuss, *supra* note 216, at 842-43.

the territoriality principle will probably enhance the willingness of courts to resort to the ALI Principles for guidance.[252] Yet, the adoption of a more traditional territorial approach came with the sacrifice of lack of clarity in cases of ubiquitous infringement, as the current version of section 321(1) only gives examples instead of definite rules.

An earlier, less territorial version of the ALI Principles endorsed a so-called "cascading" approach.[253] Its final forbearance is regrettable as the adoption of a decision cascade most likely would have entailed more definite results. Armed with a decision cascade, courts would no longer have the opportunity to make deliberate choices, but would have to follow step-by-step the connecting factors of the decision cascade. Decision cascades are quite common in general choice-of-law regimes.[254] On the basis of section 321 (1) of the ALI Principles, a decision cascade for cases of ubiquitous infringement could read as follows:

> (1) When the alleged infringing activity is ubiquitous and the laws of multiple States are pleaded, the court shall apply to the issues of existence, validity, duration, attributes, and infringement of intellectual property rights and remedies for their infringement,
>
> (a) the law of the common habitual residence of the parties if the parties had their habitual residence in the same State at the time of the infringement and if at least one of them still lives in that State; otherwise
>
> (b) the law of the State where a pre-existing relationship between the parties was centered if that relationship is closely connected with the infringement; alternatively
>
> (c) the law(s) of the State(s) towards which the parties primarily directed their activities.

In the Bill-Françoise hypothetical, a court applying the proposed provision to the case would decide the case as follows: Françoise lives in

---

252. *Id.*

253. *See* ALI, *supra* note 13, at 250.

254. On an international level, see Convention on the Law Applicable to Products Liability, art. 3-4, May 4, 1971, 37 Rabels Zeitsschrift für ausländiches und internationales Privatrecht 594. and to some extent Rome II, *supra* note 171, art. 3. On a national level, see, for example, Einführungsgesetz zum Bürgerlichen Gesetzbuche [EGBGB][Introductory Law to the Civil Code] Sept. 21, 1994, Bundesgesetzblatt [BGBl. I] III/FNA 400-1, as amended, art. 40.

Belgium, Bill resides in Great Britain, thus they do not have a common habitual residence and the first section of the proposed rule would not apply. There is also no evidence of any prior legal relationship between Bill and Françoise (e.g., a licensing agreement) that would be closely connected to the case. Consequently, the second section can also be skipped. Francoise, however, designated her work explicitly for the French market, which would lead to application of French law according to the third section of the proposed rule. If Francoise had directed her work at multiple countries, then the primary country would be the country of first publication or largest market size (in the rare cases of simultaneous publication).This solution seems perfectly in tune with general practices in copyright law as well as with common sense. The advantage of a decision cascade is that it provides courts with clear directions to determine the applicable law, which will most likely lead to more harmonized decisions on an international level. Any deviation from the traditional principle of territoriality seems justifiable given the limited applicability of the provision to cases of ubiquitous infringement where territorial consistency is no longer sustainable.

IV. CONCLUSION

The rapid technological progress that characterizes our current times has left the law struggling to keep pace. The recent establishment of the Internet as an integral part of our daily lives has had sweeping consequences for virtually all fields of law. The ease with which copyrighted material can be uploaded and made instantaneously accessible to a global audience requires a revolutionary rewrite of current copyright protection mechanisms. The existing jurisprudence in cases of copyright violations that span several countries is disparate and presents no uniformity on the method to determine the applicable law.

Three diverging methods can be extracted from the cases discussed in this Article: first, the denial of foreign copyright protection on the grounds of *forum non conveniens* and an implied application of the *lex fori* (in the Ninth Circuit in the United States and the traditional approach in Great Britain); second, application of the *lex loci delicti* (in the Second Circuit in the United States, in dicta in a recent decision by a British court, and in a recent French court decision); third, application of the *lex protectionis* (in German courts and in some new Belgian legislation). Yet, none of these countries have presented a convincing solution to the question of which law should be applied to cases of internet copyright infringement that implicate multiple countries' laws.

This Article has shown that the poor and inconsistent application of these choice-of-law rules has added further confusion and legal uncertainty. Therefore, the proper application of conflict-of-laws theory is essential to a satisfactory resolution of internet copyright infringement cases. This requires cross-fertilization of two once independently operating areas of law–intellectual property and conflict of laws. I predict that this will lead to the emergence of a novel interdisciplinary field that will require new collaborations between legal scholars of both camps.

Both the Rome II statute and the ALI Principles may be considered the first developments of such an emerging collaboration. Indeed, the application of either instrument may overcome the previous divergence in choice-of-law rules since they both have settled upon application of the *lex protectionis*. Yet, the fundamental issue of which law to choose when copyright infringement has occurred in several countries remains mostly unresolved. This Article's evaluation of the ALI Principles and the Rome II statute has revealed that the ALI Principles are better suited for the resolution of complex cases of copyright infringement on the Internet. Only the ALI Principles explicitly acknowledge the challenge that internet copyright infringement cases may present to traditional choice-of-law approaches.

Unfortunately, the solution proposed by the ALI Principles is not explicit enough to ensure a uniform approach to conflict of laws in internet copyright infringement cases. The ALI Principles list possible decision criteria without providing clear guidance how to proceed. As a result, the risk of continued confusion and lack of coherence in future jurisprudence is foreseeable.

Despite these more technical concerns, international harmonization of conflict-of-laws rules that target internet copyright infringement cases is a crucial step towards overcoming the limitations of more traditional territorial copyright law. Only the future will show how successful these recent attempts of harmonization will be. It is my hope that this Article provides valuable theoretical and practical guidance for a smooth transition from the current inconsistent and scattered national jurisprudence to an era of more consistent and comprehensive international copyright protection.

THINGS ARE WORSE THAN WE THINK: TRADEMARK DEFENSES IN A "FORMALIST" AGE

By Michael Grynberg†

TABLE OF CONTENTS

---

† Assistant Professor of Law, Oklahoma City University School of Law. My thanks to Kelly Baldrate, Michael Gibson, Art LeFrancois, Bill McGeveran, Rebecca Tushnet, and Deborah Tussey for their helpful comments and to Oklahoma City University for the summer grant that supported my research. Earlier drafts were presented at the 2008 Works in Progress Intellectual Property (WIPIP) Colloquium at Tulane University Law School and a faculty colloquium at Oklahoma City University.

To lose firm ground for once! To float! To err! To be mad!—that was part of the paradise and debauchery of former ages, whereas our bliss is like that of the shipwrecked man who has climbed ashore and is standing with both feet on the firm old earth—marveling because it does not bob up and down.

—Friedrich Nietzsche[1]

We are all textualists now.

—Hon. Marjorie Rendell[2]

---

1. FRIEDRICH NIETZSCHE, THE GAY SCIENCE 60 (Bernard Williams ed., Josefine Nauckhoff & Adrian Del Caro trans., Cambridge Univ. Press 2001) (1888).

2. Marjorie O. Rendell, *2003—A Year of Discovery: Cybergenics and Plain Meaning in Bankruptcy Cases*, 49 VILL. L. REV. 887, 887 (2004); *see also* Jonathan R. Siegel, *Textualism and Contextualism in Administrative Law*, 78 B.U. L. REV. 1023, 1057 (1998) ("In a significant sense, we are all textualists now. The days when lawyers could 'routinely . . . make no distinction between words in the text of a statute and words in its legislative history' are surely over." (quoting Antonin Scalia, *Common-Law Courts in a Civil-Law System: The Role of United States Federal Courts in Interpreting the Constitution and Laws*, in A MATTER OF INTERPRETATION: FEDERAL COURTS AND THE LAW 3, 31

I. INTRODUCTION

Numerous articles decry the expansion of trademark law. This Article assumes the premise that these critiques are valid and asks what courts can do in response. The answer may be, not much. The "common law" practices that expanded trademark's scope are not up to the task of creating adequate countervailing defenses.

It is by now commonplace to observe that trademark's domain grew considerably in the last century.[3] A once-limited remedy designed to police the false "passing off" of goods has morphed into a broad right capable of targeting mere references to popular brands.[4] Critics warn that trademark's expansion threatens competition,[5] stifles speech,[6] and creates "property" rights where none are deserved.[7] The commentary also offers a range of potential reforms. Some suggest that courts should require plaintiffs to prove that defendants engaged in a "trademark use" before the court will impose liability.[8] Others argue that courts should weigh the ef-

---

(Amy Gutmann ed., 1997)) (footnote omitted)).

3. *See, e.g.*, Robert G. Bone, *Hunting Goodwill: A History of the Concept of Goodwill in Trademark Law*, 86 B.U. L. REV. 547, 592-615 (2006) (outlining expansion); Jessica Litman, *Breakfast with Batman: The Public Interest in the Advertising Age*, 108 YALE L.J. 1717, 1722 (1999); J. Thomas McCarthy, *Lanham Act § 43(A): The Sleeping Giant is Now Wide Awake*, 59 LAW & CONTEMP. PROBS. 45, 46 (1996) ("In a half-century, section 43(a) has undergone an amazing transformation at the hands of the federal judiciary. Section 43(a) has risen from obscurity as a largely ignored subsection of the Trade Registration Act in 1945 to today's unrivaled legal instrument to combat unfair competition.").

4. *See, e.g.*, Mut. of Omaha Ins. Co. v. Novak, 836 F.2d 397 (8th Cir. 1987) (affirming a ruling that a T-shirt with the phrase "Mutant of Omaha" infringed "Mutual of Omaha" mark where a survey found that "ten percent of all the persons surveyed thought that Mutual 'goes along' with" defendant's use); Litman, *supra* note 3, at 1722 (noting enforcement of trademark claims concerning "confusion about the possibility of sponsorship or acquiescence").

5. Glynn S. Lunney, Jr., *Trademark Monopolies*, 48 EMORY L.J. 367, 485-86 (1999) (criticizing the "overbroad, ill-considered legal regime that serves simply to enrich certain trademark owners at the expense of consumers, the market's competitive structure, and the public interest more generally").

6. *See, e.g.*, Lisa P. Ramsey, *Descriptive Trademarks and the First Amendment*, 70 TENN. L. REV. 1095, 1101 (2003) (arguing that protection of descriptive trademarks fails the *Central Hudson* commercial speech test).

7. *See, e.g.*, Mark A. Lemley, *The Modern Lanham Act and the Death of Common Sense*, 108 YALE L.J. 1687, 1693 (1999) (locating a tendency of courts to treat trademarks as assets in "a broader trend towards 'propertizing' intellectual property"); Lunney, *supra* note 5, at 372 (observing rise of " 'property mania'—the belief that expanded trademark protection was necessarily desirable so long as the result could be characterized as 'property.' ").

8. *See* Margreth Barrett, *Internet Trademark Suits and the Demise of "Trademark*

fect of challenged practices on consumer search costs,[9] deemphasize the protection of a trademark holder's goodwill as a purpose of trademark law,[10] or apply stricter First Amendment scrutiny to trademark law.[11]

The underlying premise of many reform proposals is the belief that courts have the ability to implement these visions by creating new trademark defenses,[12] be it with new safe harbors, new doctrines, or modified defenses.[13] In their attack on the trademark use doctrine, for example, Graeme Dinwoodie and Mark Janis express confidence that courts have authority to craft more targeted defenses that will protect trademark defendants who engage in socially beneficial activities.[14] Their sparring

---

Use", 39 U.C. DAVIS L. REV. 371 (2006); Stacey L. Dogan & Mark A. Lemley, *Grounding Trademark Law Through Trademark Use*, 92 IOWA L. REV. 1669 (2007); Uli Widmaier, *Use, Liability, and the Structure of Trademark Law*, 33 HOFSTRA L. REV. 603 (2004).

9. *See, e.g.*, Stacey L. Dogan & Mark A. Lemley, *Trademarks and Consumer Search Costs on the Internet*, 41 HOUS. L. REV. 777 (2004).

10. *See, e.g.*, Bone, *supra* note 3, at 616–22; Michael Grynberg, *Trademark Litigation as Consumer Conflict*, 83 N.Y.U. L. REV. 60, 116-17 (2008).

11. *See, e.g.*, Ramsey, *supra* note 6, at 1176 ("The First Amendment requires the government to revise the trademark laws to prevent registration and enforcement of exclusive rights in descriptive terms."); *cf.* Rebecca Tushnet, *Trademark Law as Commercial Speech Regulation*, 58 S.C. L. REV. 737, 755 (2007) ("Taking modern First Amendment doctrine seriously would have significant effects on the Lanham Act, affecting everything from the standard of proof to the definition of what counts as misleading.").

12. Unless otherwise noted, when I use the term "defense," I mean it broadly to encompass any doctrine that a defendant may invoke to defeat a trademark claim notwithstanding the plaintiff's ability to establish that a likelihood of confusion exists.

13. *See, e.g.*, Graeme B. Dinwoodie, *Developing Defenses in Trademark Law*, 13 LEWIS & CLARK L. REV. 99, 112 (2009) [hereinafter Dinwoodie, *Developing Defenses*] ("[T]rademark law would be well-served by the development of real defenses more generally."); William McGeveran, *Rethinking Trademark Fair Use*, 94 IOWA L. REV. 49, 115-121 (2008) (proposing safe harbors); Mark P. McKenna, *Trademark Use and the Problem of Source*, 2009 U. ILL. L. REV. (forthcoming 2009) (manuscript at 78), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1088479 ("The most obvious way to curb further expansion of trademark rights is to recognize additional doctrines, like functionality, that are outcome determinative without regard to consumer understanding. Some existing defenses could be made more independent simply by ceasing to condition the defenses on lack of confusion.") (footnote omitted).

14. Graeme B. Dinwoodie & Mark D. Janis, *Lessons from the Trademark Use Debate*, 92 IOWA L. REV. 1703, 1708-09 (2007) [hereinafter Dinwoodie & Janis, *Lessons*] ("[O]ur contextual approach contemplates that courts will continue to develop defenses as they are called upon to balance confusion-avoidance values against other values in new contexts. And we believe that the trademark statute provides them plenty of room do so.") (footnote omitted); *see also* Graeme B. Dinwoodie & Mark D. Janis, *Confusion Over Use: Contextualism in Trademark Law*, 92 IOWA L. REV. 1597, 1616 (2007) [hereinafter Dinwoodie & Janis, *Contextualism*] ("U.S. trademark law has long recognized

partners and other commentators agree.[15]

This Article takes issue with this faith and argues that efforts to reform trademark law with new defenses lack a firm basis. At first glance, this claim may seem inconsistent with the history of trademark law. Courts nursed trademark's expansion by acting in the common law tradition, crafting judicial rules to respond to perceived needs. Even after the Lanham Act[16] filled the federal common law void left by *Erie Railroad Co. v. Tompkins*,[17] judicial elaboration of federal trademark law frequently left statutory text behind. Congress acquiesced by amending the law to conform to judicial interpretation of the original statute. Today's Lanham Act permits—even if it does not compel—a robust federal trademark law.[18]

Just as trademark liability has roots in the common law, so too do trademark defenses. Many of these defenses are now codified.[19] If the pendulum has swung too widely with respect to trademark rights, might courts employ similar "common law" decision making to create defenses to circumscribe trademark liability?

Perhaps not. We live in a different world, one in which judges are less confident about crafting rules to supplement statutes, and the Supreme Court routinely reminds litigants that the days of implied causes of action are over.[20] This reticence is manifest in the Court's recent resistance to assertions of trademark rights beyond the express confines of the Lanham Act.[21] If these outcomes bespeak a more formalist, textualist approach to

---

extra-statutory defenses to statutory causes of action.").

15. Dogan & Lemley, *supra* note 8, at 1685 ("The fact that Congress has codified explicit exclusions for some of these categories does not foreclose courts from recognizing others, particularly those that have long been implicit in trademark law. Indeed, most of the exclusions and defenses in the Lanham Act—including descriptive fair use—began as common-law doctrines."); McGeveran, *supra* note 13, at 121 ("Fortunately, common-law reasoning is alive and well in trademark law. Indeed, the entire structure of likelihood of confusion reasoning is extra-statutory. Just as courts created nominative fair use and First Amendment balancing as common law, they could establish safe harbors."); McKenna, *supra* note 13 (manuscript at 78) ("Courts could determine, for example, that comparative advertising has social benefit and is not infringement *even if it causes some confusion*.") (emphasis added).

16. 15 U.S.C. §§ 1051–1141n (2006).

17. 304 U.S. 64 (1938).

18. *See infra* Section II.A.4.

19. 15 U.S.C. § 1115(b) (2006).

20. *See, e.g.*, Stoneridge Inv. Partners, L.L.C. v. Scientific-Atlanta, Inc., 128 S. Ct. 761, 772 (2008) ("[I]t is settled that there is an implied cause of action only if the underlying statute can be interpreted to disclose the intent to create one . . . ."); *see infra* Section III.A.

21. *See, e.g.*, Moseley v. V Secret Catalogue, Inc., 537 U.S. 418, 433-34 (2003)

trademark law, so much the worse for defensive innovations, which lack the broadly worded statutory language that continues to nourish expansive liability. Put another way, the swinging pendulum may have slowed trademark's spread, but past gains are safe.

Trademark defenses already enjoy an uneasy status in federal law. Trademark liability rests on broadly worded text. The Lanham Act contains two causes of action for trademark infringement—found in sections 32 and 43(a) of the statute. The first applies to infringement of registered marks,[22] the second sweeps more broadly to reach any use of a word, symbol, or device that is likely to cause confusion.[23] By contrast, trademark defenses are less expansive.[24] The Act's statutory defenses only expressly apply to the cause of action for *registered* marks.[25] While the application of these defenses to section 43(a) is arguably textually legitimate, either as an exercise of federal common law or as statutory construction, there is little room for courts to go further.[26]

What does this mean for the future of trademark law? If judges have only a limited ability to create trademark defenses, then broadened trademark defenses may require congressional action or reliance upon extrinsic

---

(holding that the federal dilution statute required proof of actual dilution and not merely a likelihood of dilution). Congress negated this holding by passing the Trademark Dilution Revision Act of 2006, Pub. L. No. 109-312, 120 Stat. 1730 (amending scattered sections of title 15 of the United States Code); Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23, 32-37 (2003) (holding that the Lanham Act's prohibition of false designations of origin do not prohibit uncredited copying of another's work); *see generally infra* Section III.B.

22. 15 U.S.C. § 1114 (2006) (proscribing the "use in commerce [of] any reproduction, counterfeit, copy, or colorable imitation of a registered mark in connection with the sale, offering for sale, distribution, or advertising of any goods or services on or in connection with which such use is likely to cause confusion, or to cause mistake, or to deceive.").

23. 15 U.S.C. § 1125(a) (2006) (prohibiting use of "any word, term, name, symbol, or device, or . . . any false designation of origin, false or misleading description of fact, or false or misleading representation of fact" that "is likely to cause confusion, or to cause mistake, or to deceive as to the affiliation, connection, or association of such person with another person, or as to the origin, sponsorship, or approval of his or her goods, services, or commercial activities by another person").

24. *See infra* Section II.B.

25. 15 U.S.C. § 1115(b) (2006) (setting forth defenses to incontestable registered marks); *id.* § 1115(a) (providing that defenses to actions involving incontestable marks apply to those involving any registered mark).

26. *See infra* Section IV.A.2; 5 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 27:19 (4th ed. 2009) ("[T]he statutory 'defenses' in a § 43(a) case are merely guidelines to ascertain the federal common law substantive 'defenses' to a § 43(a) claim.").

sources of law, such as the First Amendment. Beyond that, the best hope may be for courts to roll back the expansion that they have paced. Congress may have given the courts a lever with which to control trademark's scope, but it is a lever that operates on liability, not defenses. And it is in the open wording of the Lanham Act's liability-creating provisions that judicial flexibility, to the extent that it exists, is to be found.

Part II outlines the problem created by trademark's expansion and the limited role that current doctrine leaves for defenses. Part III identifies trademark's "common law" problem and explains why the judicial style that spurred trademark's advance is not necessarily available to curtail it. Part IV discusses the consequences for future trademark defenses. Part V proffers potential solutions to the defense dilemma that are rooted in the text of the Lanham Act, while urging continued resistance of trademark's expansion. My argument nonetheless concludes on a pessimistic note. The doctrinal realities of modern trademark law make reform efforts based on judicial action unlikely to succeed.

II. TRADEMARK'S EXPANSION AND THE LIMITS OF TRADEMARK DEFENSES

This Part explains the need for further development of trademark defenses. While broad and flexible doctrines set the scope of trademark liability, defenses to infringement are comparatively narrow and rigid. Consequently, they are often ill-equipped to act as a check against ambitious plaintiffs.

A. Trademark's Expanding Scope

Trademarks and servicemarks perform the basic function of allowing sellers to brand their goods and services.[27] The Lanham Act protects marks by providing causes of action when a junior user's mark is likely to cause confusion with that of the senior user.[28]

---

27. 15 U.S.C. § 1127 (2006) (defining trademarks and servicemarks). Unless otherwise noted, this Article uses the term "trademarks" to encompass both trademarks and servicemarks.

28. *See* 15 U.S.C. § 1114 (establishing liability for conduct that "is likely to cause confusion, or to cause mistake, or to deceive" with respect to registered marks); § 1125(a)(1)(A) (creating liability under "common law" trademark infringement action for conduct that "is likely to cause confusion, or to cause mistake, or to deceive as to the affiliation, connection, or association of such person with another person, or as to the origin, sponsorship, or approval of his or her goods, services, or commercial activities by another person"). Though state laws may, and sometimes do, grant broader rights than the federal statute (subject to preemption principles), 3 MCCARTHY, *supra* note 26, § 22:2,

A trademark plaintiff must meet several conditions for a successful claim. First, she must possess a valid trademark. This requirement encompasses the questions of mark ownership—as determined by prior use or registration[29]—and mark eligibility. The mark must be able to identify and distinguish goods and services in the marketplace, while avoiding various exclusions to trademark status.[30]

If the senior user has a valid mark, she must establish that the defendant is using it (or something similar) in a manner likely to cause consumer confusion,[31] which raises the difficult questions of what consumers must be confused about, when consumer confusion is measured, and how do we measure it?

Today, plaintiffs have an easier time clearing the hurdles described above. While courts have expanded the potential scope of trademark claims, they have been less diligent about erecting barriers to potential plaintiff overreaching.

The story of trademark's growth is oft-told and will only be summarized briefly in this section.[32] The key point for present purposes is that this growth was largely a judicial creation.[33] Judges gave expansive interpretations to seemingly restrictive statutory text, and Congress gave its blessing to the results.

1. What May Be a Trademark?

Federal trademark protection was once limited to "technical" trademarks, which encompassed inherently distinctive marks and excluded trade names or dress regardless of the acquisition of secondary meaning (i.e., association by consumers of the identifying device with a single source).[34] Today the Lanham Act broadly defines trademark to include

---

today most state law unfair competition actions parallel their federal counterpart. 4 *id.* § 23:1.50 ("Most courts, in analyzing a claim of infringement based on both federal and state law, will apply to both a single analysis of the likelihood of confusion issue.").

29. 15 U.S.C. § 1057(c) (2006) (declaring registration of a mark to give nationwide priority subject to rights of prior users, applicants, or foreign registrants meeting certain conditions).

30. 15 U.S.C. § 1127 (declaring that trademarks and servicemarks "identify and distinguish" goods and services); *id.* § 1052 (setting forth registration requirements and bars to registration).

31. 15 U.S.C. §§ 1114, 1125(a)(1).

32. For some sources telling the tale, see *supra* note 3.

33. *See, e.g.*, U.S. Trademark Ass'n Trademark Review Comm'n, *Report and Recommendations to USTA President and Board of Directors*, 77 TRADEMARK REP. 375, 376 (1987) [hereinafter USTA Report] ("In the 1970s the courts transformed [section 43(a)] into a potent, far-reaching, commercial Bill of Rights for the honest businessman.").

34. 1 MCCARTHY, *supra* note 26, § 4:5 ("Under archaic usage, marks that were not

"any word, name, symbol, or device, or any combination thereof" used to distinguish goods in the marketplace.[35] From this point of departure, the Supreme Court approved trademark protection for colors[36] and unregistered, but distinctive, trade dress.[37] In the latter case, the Court did so de-

---

inherently distinctive were not protected as 'technical trademarks,' but were protected as 'trade names' under the law of 'unfair competition' upon proof of secondary meaning."); *id.* § 4:12 ("The Lanham Act of 1946 integrated the two types of common law marks (technical trademarks and trade names), calling both types 'trademarks' and treating them in essentially the same manner."); *id.* § 8:1 (noting that law of unfair competition encompassed trade dress and "[a]s with archaic 'trade names,' trade dress protected under the law of 'unfair competition' always required proof of secondary meaning"). Early federal trademark statutes restricted registration to these technical trademarks. *Id.* § 5:3; 1 WILLIAM D. SHOEMAKER, TRADE-MARKS: A TREATISE ON THE SUBJECT OF TRADE-MARKS WITH PARTICULAR REFERENCE TO THE LAWS RELATING TO REGISTRATION THEREOF 236 (1931) ("Physical characteristics of an article, its appearance, style or dress-up or features of containers or wrappers cannot be subject of exclusive appropriation . . . .").

35. 15 U.S.C. § 1127; *cf.* RESTATEMENT (THIRD) OF UNFAIR COMPETITION, § 16 cmt. a (1995) ("With the abandonment of the distinction between technical 'trademarks' and other indicia of source, the protection of distinctive packaging and product design has been incorporated into the general law of trademarks.").

36. Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159 (1995). The Court dismissed the possibility that permitting colors to be trademarked would give rise to too many suits or inhibit competition by depleting the supply of available colors in the marketplace. *Id.* at 167-70. To be sure, *Qualitex* nodded at the concern by suggesting that secondary meaning was required for color to function as a mark. *Id.* at 163 ("We cannot find in the basic objectives of trademark law any obvious theoretical objection to the use of color alone as a trademark, where that color has attained 'secondary meaning' and therefore identifies and distinguishes a particular brand (and thus indicates its 'source').."). The Court later clarified that it did intend to so hold. Wal-Mart Stores, Inc. v. Samara Bros., 529 U.S. 205, 211 (2000) ("Indeed, with respect to at least one category of mark-colors we have held that no mark can ever be inherently distinctive." (citing *Qualitex*, 514 U.S. at 162-63)). *Compare* 1 SHOEMAKER, *supra* note 34, at 163 ("Whether mere color can constitute a valid trade-mark may admit of doubt. Doubtless it may be, if it be impressed in a particular design, as a circle, square, triangle, a cross or a star. But the authorities do not go further than this."), *with* NORMAN F. HESSELTINE, A DIGEST OF THE LAW OF TRADE-MARKS & UNFAIR TRADE 81, 81-82 (1906).

37. "Trade dress" is undefined by the Lanham Act, but is the term used to encompass the trademark functions performed by product packaging and/or design—i.e., the total marketplace presentation. RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 16 cmt. a (1995) ("The term 'trade dress' is often used to describe the overall appearance or image of goods or services as offered for sale in the marketplace."). The question at issue in *Two Pesos, Inc. v. Taco Cabana, Inc.* was whether trade dress may be inherently distinctive, that is, be treated as a trademark without evidence that consumers had come to associate the trade dress with a single source. The alternative was first to require evidence that the dress had acquired "secondary meaning" with the consuming public. Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763, 765 (1992). The Court held that secondary meaning was not required. *Id.* at 776. The Court later held, however, that secondary meaning was required if protection is sought for an allegedly distinctive product design

spite a strong textual argument that protection of unregistered trade dress fell outside the Lanham Act's original scope.[38] Justice Stevens, who made the argument, nonetheless concurred in the result, in part because the then-recent amendment of the statute indicated Congress's support of protection.[39]

2. *About What Must Consumers Be Confused?*

Justice Stevens based his skepticism of strong trade dress protection on section 43(a)'s original text, which restricted liability to defendants who "affix, apply, or annex, or use" a "false designation of *origin*."[40] The statutory text strongly implied that "origin" meant *geographic* origin insofar as it provided a cause of action "by any person doing business *in the locality falsely indicated as that of origin* or the region in which said locality is situated."[41] And indeed, for a time, practitioners perceived the scope of the cause of action as narrow.[42] Section 43(a) also had prohibited "any false description or representation," which lacked obvious applicability to trade dress.[43]

(as opposed to packaging). Wal-Mart Stores, Inc., 529 U.S. 205 (2000).

38. *Two Pesos*, 505 U.S. at 777-79 (Stevens, J., concurring in the judgment); *id.* at 781 ("Even though the lower courts' expansion of the categories contained in § 43(a) is unsupported by the text of the Act, I am persuaded that it is consistent with the general purposes of the Act."). While somewhat dismissive of Justice Stevens's argument in his *Two Pesos* concurrence, *id.* at 776 (Scalia, J., concurring), Justice Scalia acknowledged its force in his later opinion for the Court in *Dastar*, noting that "a case can be made that a proper reading of § 43(a), as originally enacted, would treat the word 'origin' as referring only 'to the geographic location in which the goods originated,' " *Dastar*, 539 U.S. at 29 (quoting *Two Pesos*, 505 U.S. at 777 (Stevens, J., concurring in the judgment)).

39. *Two Pesos*, 505 U.S. at 776 (Stevens, J., concurring in the judgment) ("I agree with this transformation, even though it marks a departure from the original text, because it is consistent with the purposes of the statute and has recently been endorsed by Congress."). Specifically, Congress added language that "make[s] explicit that the provision prohibits 'any word, term, name, symbol, or device, or any combination thereof' " that is likely to cause confusion. *Id.* at 783 (quoting 15 U.S.C. § 1125(a)).

40. Trademark (Lanham) Act of 1946, ch. 540, § 43, 60 Stat. 427, 441 (current version at 15 U.S.C. § 1125 (2006) (emphasis added).

41. *Id.* (emphasis added); *see also* McCarthy, *supra* note 3, at 47-48 (observing that a restrictive view of § 43(a) was "conventional wisdom in 1946" and that the "future expansive possibilities of section 43(a) were only vaguely perceived at that time").

42. McCarthy, *supra* note 3, at 52 ("In 1956, Judge Clark of the Second Circuit remarked of section 43(a) that 'the bar has not yet realized the potential impact of this statutory provision.' ").

43. Lanham Act § 43, 60 Stat. at 441 (current version at 15 U.S.C. § 1125 (2006)). That is, a trade dress does not seem to be an affixed "description or representation," certainly not the sort of trade dress at issue in *Two Pesos*. *See Two Pesos*, 505 U.S. at 778 (Stevens, J., concurring in the judgment) (arguing that the language served only to police

Similarly, the original cause of action for infringement of a registered trademark only extended to the use of a "reproduction, counterfeit, copy, or colorable imitation of any registered mark" where such use is likely to "cause confusion or mistake or to deceive *purchasers* as to the *source of origin of such goods or services*."[44] Subsequently, Congress dropped the purchaser limitation and origin language, provoking interpretive debate as to the 1962 amendment's significance.[45]

From these statutory roots, federal liability blossomed. Courts extended the section 43(a) cause of action to reach beyond designations of origin to encompass situations pertaining to confusion of source[46] or spon-

---

false advertising and the common law tort of passing off). *But see id.* at 787 (Thomas, J., concurring in the judgment) (arguing that the language encompasses trade dress protection).

44. Lanham Act § 32, 60 Stat. at 437 (current version at 15 U.S.C. § 1114 (2006)) (emphases added).

45. Act of Oct. 9, 1962, Pub. L. No. 87-772, sec. 17, § 32, 76 Stat. 769, 773 (current version at 15 U.S.C. § 1114 (2006)). The USTA Report noted that courts took the amendment to mean more than it did. "The change was explained, innocently enough, as parallel to a similar change being made in Section 2(d)," which pertained to a registration bar for marks that were likely to cause confusion and was amended to by deleting the "purchasers" term "to make it clear that the provision related to potential as well as actual purchasers." USTA Report, *supra* note 33, at 378. Instead, "a number of courts have viewed the deletion as evidence of Congressional intent to broaden the test for likelihood of confusion. Now, they say, the Act is designed to prohibit confusion of any kind, not merely of purchasers or customers nor as to source of origin." *Id. Compare* Checkpoint Sys., Inc. v. Check Point Software Techs., Inc., 269 F.3d 270, 295 (3d Cir. 2001) ("[W]e agree with the view that Congress's amendment of the Lanham Act in 1962 expanded trademark protection to include instances in which a mark creates initial interest confusion."), *with* Elec. Design & Sales, Inc. v. Elec. Data Sys. Corp., 954 F.2d 713, 716 (Fed. Cir. 1992) ("We do not construe this deletion to suggest, much less compel, that purchaser confusion is no longer the primary focus of the inquiry.").

46. McCarthy, *supra* note 3, at 51-52; *id.* at 59 ("While this expansion of the word source was criticized as unwise, by the early 1980s it had become a firmly embedded reality.") (footnotes omitted). Professor McCarthy explains:

> The phrase "false designation of origin" was thought to be limited to false advertising of geographic origin. The first expansion of the meaning of "origin" to include origin of source, sponsorship or affiliation in the classic trademark sense, came in 1963. In that year, the U.S. Court of Appeals for the Sixth Circuit held that "origin" did not refer only to geographic origin but also "to origin of source of manufacture." This seemingly simple new spin put on the word "origin" raised the curtain on a whole new chapter in federal unfair competition law. It heralded the beginning of a new dimension of section 43(a) as a vehicle to assert in federal court a traditional case of infringement of an unregistered mark, name, or trade dress.

Id. at 58 (footnotes omitted).

sorship.[47] Other extensions stretched the concept of sponsorship from a form of endorsement or guarantee of quality to the mere permission to engage in a particular use.[48] As a result, trademark holders may be able to take control of merchandising markets in which a mark—e.g., a sports team logo on a baseball cap—is not serving as a source designation, but is rather the product itself.[49] Other courts recognized claims based on purported likelihood of consumer confusion that foreshadowed the later enacted federal dilution statute.[50]

3. *When Do We Measure Confusion?*

Courts now assess likely confusion at times other than the point of purchase. For instance, courts adopting the theory of initial interest confusion assign liability for confusion even if it is dispelled before the point of sale.[51] Courts have applied the doctrine to a range of activities deemed to have "diverted" a consumer's attention even if she knows what she is buying when she pays her money.[52] In other cases, the trademark cause of action has proven robust enough to include confusion of non-purchasers who view a product after purchase.[53]

---

47. 4 MCCARTHY, *supra* note 26, § 24:6. Professor McCarthy writes that section 43(a) was originally seen as a minor provision that might be helpful in false advertising cases. *Id.* § 27:7.

48. *See, e.g.*, *supra* note 4.

49. In such cases, the purchaser of the cap is attracted to the trademarked logo not because he believes that the mark owner is the physical source of the hat, but because he wants a hat with that particular logo. In this scenario, the logo is not a mark in the traditional sense, but rather an indispensable feature of the product. Many courts allow trademark holders to use the Lanham Act to restrict competition in the merchandising market. *See, e.g.*, Boston Prof'l Hockey Ass'n v. Dallas Cap & Emblem Mfg., Inc., 510 F.2d 1004 (5th Cir. 1975); *see generally* Stacey L. Dogan & Mark A. Lemley, *The Merchandising Right: Fragile Theory or Fait Accompli?*, 54 EMORY L.J. 461, 472-78 (2005) (surveying judicial treatment of the merchandising right).

50. Under the current federal dilution statute, dilution constitutes two harms: dilution by blurring and dilution by tarnishment. Dilution by blurring is "association arising from the similarity between a mark or trade name and a famous mark that impairs the distinctiveness of the famous mark." 15 U.S.C. § 1125(c) (2006). Dilution by tarnishment is "association arising from the similarity between a mark or trade name and a famous mark that harms the reputation of the famous mark." *Id.* For an example of tarnishment reasoning appearing in a likelihood of confusion case, see, e.g., Dallas Cowboys Cheerleaders, Inc. v. Pussycat Cinema, Ltd., 604 F.2d 200, 205 (2d Cir. 1979) (upholding district court's award of a preliminary injunction against adult film depicting characters with uniforms similar to that of professional team's cheerleaders).

51. *See, e.g.*, Grotrian v. Steinway & Sons, 523 F.2d 1331, 1342 (2d Cir. 1975).

52. *See infra* notes 102-104 and accompanying text. *Compare* McCarthy, *supra* note 3, at 50 (noting that section 43(a) does not cover "bait-and-switch selling tactics").

53. *See, e.g.*, Hermès Int'l v. Lederer de Paris Fifth Ave., Inc., 219 F.3d 104, 108

4. Expansive Trademark and the Lanham Act

The judicial branch's leading role in trademark law's expansion is problematic.[54] Though the Lanham Act contains open provisions, it does not grant courts common law authority over unfair competition generally.[55] Yet many interpretations discussed above were, at least arguably, hostile to the statute's text and more consistent with common lawmaking than statutory interpretation.[56]

To some extent, this is unsurprising. The Lanham Act—like the law of trademarks more generally—leaves much to the judicial imagination. Most fundamentally, the basic fact question of whether consumers are likely to be confused is a murky one.[57] Trademark's roots, moreover, are

---

(2d Cir. 2000) (condemning sales of knockoff products "for the purpose of acquiring the prestige gained by displaying what many visitors at the customers' homes would regard as a prestigious article" (quoting Mastercrafters Clock & Radio Co. v. Vacheron & Constantin-LeCoultre Watches, Inc., 221 F.2d 464, 466 (2d Cir. 1955))).

54. A point not lost on Justice Stevens in *Two Pesos*. *See supra* note 39.

55. *See* S. REP. NO. 79-1333 (1946), *as reprinted in* 1946 U.S.C.C.A.N. 1274, 1276-77 (observing that trademark rights were once largely based in the common law, but Supreme Court's conclusion that federal common law does not exist, coupled with rise of national markets, necessitated statutes that would creates national rights).

56. The expansion of the original section 43(a)'s provision regarding "origin" is one example. *See supra* notes 38-42 and accompanying text. For an especially tortured interpretation of "likelihood of confusion," see Boston Prof'l Hockey Ass'n v. Dallas Cap & Emblem Mfg., Inc., 510 F.2d 1004, 1012 (5th Cir. 1975). The court noted, in a case involving sports team emblems:

> It can be said that the public buyer knew that the emblems portrayed the teams' symbols. Thus, it can be argued, the buyer is not confused or deceived. This argument misplaces the purpose of the confusion requirement. The confusion or deceit requirement is met by the fact that the defendant duplicated the protected trademarks and sold them to the public knowing that the public would identify them as being the teams' trademarks. The certain knowledge of the buyer that the source and origin of the trademark symbols were in plaintiffs satisfies the requirement of the act.

Id.

57. Courts have developed multifactor tests to guide the analysis. 4 MCCARTHY, *supra* note 26, § 24:30-43 (listing factors used by various circuits). For example, the *Lapp* factors of the Third Circuit are:

> (1) the degree of similarity between the owner's mark and the alleged infringing mark; (2) the strength of the owner's mark; (3) the price of the goods and other factors indicative of the care and attention expected of consumers when making a purchase; (4) the length of time the defendant has used the mark without evidence of actual confusion arising; (5) the intent of the defendant in adopting the mark; (6) the evidence of actual confusion; (7) whether the goods, though not competing, are

in the common law,[58] and common law practices may well persist even after the passage of a somewhat comprehensive federal statute.[59] Finally, courts were operating in what has been described as a less formalist age—one in which courts were more likely to bend a statute's text in an effort to conform to Congress's perceived purposes even if those purposes were not manifest from the words of the enacted legislation.[60]

> marketed through the same channels of trade and advertised through the same media; (8) the extent to which the targets of the parties' sales efforts are the same; (9) the relationship of the goods in the minds of consumers because of the similarity of function; (10) other facts suggesting that the consuming public might expect the prior owner to manufacture a product in the defendant's market, or that he is likely to expand into that market.

Interpace Corp. v. Lapp, Inc., 721 F.2d 460, 463 (3d Cir. 1983).

The number and, at times, vagueness of these tests leave them open to manipulation by the factfinder, particularly a factor like good faith, which lacks a necessary nexus to existence of likelihood of confusion. Grynberg, *supra* note 10, at 69. In an empirical study of the various circuits' applications of the tests, Barton Beebe has found that a relatively small number of factors predominate, leaving courts to "stampede" the remainder once a determination is made based on the critical factors. Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581, 1581-82 (2006).

The elusive nature of the underlying inquiry similarly invites appellate overreaching. Reviewing courts may scrutinize the lower court's application of certain test factors in order to second guess trial-level determinations of questions of fact. *See generally* 4 MCCARTHY, *supra* note 26, § 23:73 (surveying circuit standards of appellate review on likelihood of confusion question).

58. As a 1931 treatise observed, "[s]ince a trade-mark right is a common law right, defined by the common law, the essentials of such a right are measured and analyzed by the pronouncements of the courts as to what the common law is on these points." 1 SHOEMAKER, *supra* note 34, at 1.

59. Such expansions were not inevitable. Writing after the passage of the Lanham Act, Bartholomew Diggins observed that the statute self-consciously addressed trademarks and not the law of unfair competition as a whole. Bartholomew Diggins, *The Lanham Trade-Mark Act*, 35 GEO. L.J. 147, 150, 153 (1947); *see also* McCarthy, *supra* note 3, at 50-51. But then, as now, the Lanham Act was not a masterpiece of legislative drafting; it invited judicial play at the joints, as observed at its initial passage. *See* Diggins, *supra*, at 208 ("As a statute, the Act is not well drafted and many of its provisions are ambiguous or even contradictory. Extensive litigation is almost inevitable and the courts will be faced with difficult issues of statutory construction.").

60. *See infra* Section III.A. For example, section 43(a)'s expansion was not only with respect to scope, but also remedies, as courts routinely made available remedies for infringement of *registered* marks (provided by section 35 and including profits and damages) available under section 43(a) even though section 35, by its terms, applied only to registered marks. In *Rickard v. Auto Publisher, Inc.*, 735 F.2d 450 (11th Cir. 1984), for example, the court identified no statutory basis for overlooking the plain meaning of the text. Instead, it manufactured an ambiguity based on the fact that earlier holdings had

In any event, "[i]t was the federal courts that filled section 43(a) with this new and potent content" by broadly interpreting "seemingly narrow and innocuous statutory language."[61] Small wonder then that shortly before the passage of the 1988 amendments to the act, a body of the U.S. Trademark Association could declare confidently that "under the rubric of Section 43(a), there is in every way but name only a federal common law of the major branches of the law of unfair competition."[62] The breadth of the section 43(a) cause of action, in turn, shifted the focus of trademark litigation from state to federal courts and invited further development of federal trademark law.[63]

While judicial interpretation of the Lanham Act may have once outpaced its text, today's broad cause of action has a clear statutory basis. Congress embraced the liberties taken with the original statute by rewriting the Lanham Act to conform to the courts' interpretive practices.[64] Af-

---

applied section 35 to section 43(a). *Id.* at 455. It then resolved the purported ambiguity by looking to its perceived purposes of the Act concluding that they would be best served by applying section 35 remedies to section 43(a) actions. *Id.* at 457-58.

61. McCarthy, *supra* note 3, at 45-46.

62. USTA Report, *supra* note 33, at 376; *see also* Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763, 780 (1992) (Stevens, J., concurring). Justice Stevens agreed:

> Section 43(a) is an enigma, but a very popular one. Narrowly drawn and intended to reach false designations or representations as to the geographical origin of products, the section has been widely interpreted to create, in essence, a federal law of unfair competition. . . . It has definitely eliminated a gap in unfair competition law, and its vitality is showing no signs of age.

Id. (quoting USTA Report, *supra* note 33, at 426) (ellipses in original). *But cf.* Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23, 29 (2003) ("[B]ecause of its inherently limited wording, § 43(a) can never be a federal 'codification' of the overall law of 'unfair competition,' but can apply only to certain unfair trade practices prohibited by its text.") (citation omitted) (internal quotations omitted). The USTA Report was the basis for the 1988 Lanham Act amendments. *See* S. REP. NO. 100-515, at 2 (1988), *as reprinted in* 1988 U.S.C.C.A.N. 5577, 5578.

63. McCarthy, *supra* note 3, at 74 ("Before passage of the Lanham Act, such issues were largely played out in the context of state common law. Today, the battleground is section 43(a).").

64. And it did so self consciously according to the Senate report:

> [The bill] revises Section 43(a) of the Act (15 U.S.C. 1125(a)) to codify the interpretation it has been given by the courts. Because Section 43(a) of the Act fills an important gap in federal unfair competition law, the committee expects the courts to continue to interpret the section.
>
> As written, Section 43(a) appears to deal only with false descriptions or representations and false designations of geographic origin. Since its enactment in 1946, however, it has been widely interpreted as

ter amendment in 1988, section 43(a) creates liability for use of "any word, term, name, symbol, or device, or any combination thereof" that "is likely to cause confusion, or to cause mistake, or to deceive as to the *affiliation*, *connection*, or *association* of such person with another person, or as to the origin, sponsorship, or *approval* of his or her goods, services, or commercial activities by another person."[65] As planned,[66] this change allowed broad causes of action without requiring radical stretches of the Act's text.[67]

Contemporary trademark liability standards lack clear textual limitations. For illustration, one need look no further than the cornerstone of lia-

---

> creating, in essence, a federal law of unfair competition. For example, it has been applied to cases involving the infringement of unregistered marks, violations of trade dress and certain nonfunctional configurations of goods and actionable false advertising claims.

S. REP. NO. 100-515, at 40 (1988), *as reprinted in* 1988 U.S.C.C.A.N. 5577, 5603; *see also* USTA Report, *supra* note 33, at 426 ("The Commission was reluctant to recommend any change at all [to section 43(a)]. However, to prevent judicial back-tracking . . . the Commission believes it advisable to conform the language of Section 43(a) to the expanded scope of protection applied by the courts.").

Section 43(a)'s liability provision was not the only area of congressional acquiescence. As the report noted, courts "with increased frequency [were] disregarding" the Lanham Act's text and applying the remedies available for infringement of registered marks to 43(a) violations even though "[a]s written, the remedy sections of the Lanham Act . . . apply only to violations of a registered trademark." S. REP. 100-515 at 39, 1988 U.S.C.C.A.N. at 5601-02 (citations omitted); *see also supra* note 60. Here, too, the response was to amend the Act to conform to the unfaithful judicial practice. S. REP. 100-515 at 39-40, 1988 U.S.C.C.A.N. at 5602.

65. Trademark Law Revision Act of 1988, Pub. L. 100-667, sec. 132, § 43, 102 Stat. 3935, 3946 (codified as amended at 15 U.S.C. § 1125 (2006)).

66. USTA Report, *supra* note 33, at 435-36 ("In drafting the foregoing language the Commission in no way intended to limit the continuously expanding scope of Section 43(a) as developed in forty years of decisions. We trust we have left unlimited room for the courts to expand even further this vigorous section.").

67. See, for example, *Dastar Corp. v. Twentieth Century Fox Film Corp.*, which noted:

> Under the 1946 version of the Act, § 43(a) was read as providing a cause of action for trademark infringement even where the trademark owner had not itself produced the goods sold under its mark, but had licensed others to sell under its name goods produced by them-the typical franchise arrangement. This stretching of the concept "origin of goods" is seemingly no longer needed: The 1988 amendments to § 43(a) now expressly prohibit the use of any "word, term, name, symbol, or device," or "false or misleading description of fact" that is likely to cause confusion as to "affiliation, connection, or association . . . with another person," or as to "sponsorship, or approval" of goods.

539 U.S. 23, 32 n.5 (2003) (citations omitted).

bility: the existence of a "likelihood of confusion." The Act offers neither a quantitative nor a qualitative definition of the term. That is, the text does not explain how likely confusion must be before a court may act. Courts are therefore free to find a likelihood of confusion even if consumer surveys indicate that consumers are more likely than not to avoid confusion.[68] Similarly, there is no requirement that any likely consumer confusion be material to a purchasing decision.[69]

Likewise, the Lanham Act's authorization of claims based on alleged confusion as to whether a trademark holder "approv[es]" of a use of her mark[70] provides textual support for cases in which the claimed confusion, if any, is simply over a point of law, specifically, whether the trademark holder has the right to enjoin the use of his marks even though the mark is not used to indicate a product's source or origin.[71] Because "approval" is undefined, the term is textually capable of application to such claims.

In sum, whatever one may say about trademark's past expansion as an exercise of fealty to congressional will, Congress has endorsed the result with language capable of expansive interpretation.[72] Perhaps the Lanham Act's causes of action should be read narrowly, perhaps broadly, but the modern act's text does not resolve the matter.[73]

---

68. *See* 4 MCCARTHY, *supra* note 26, § 32:188 (discussing surveys deemed probative of likelihood of confusion).

69. Materiality considerations are not wholly absent from the Lanham Act. For example, "deceptive" trademarks may not be registered, and courts look to the materiality of misrepresentations to consumers to determine whether misdescriptive marks are deceptive. *See, e.g.*, *In re* Budge Mfg. Co., 857 F.2d 773 (Fed. Cir. 1988) (applying a materiality test to determine whether a trademark is "deceptive" and therefore ineligible for registration under section 2 of Lanham Act). Similarly, section 43(a)'s cause of action for false advertising has been interpreted to include a materiality element. *See, e.g.*, Time Warner Cable, Inc. v. DIRECTV, Inc., 497 F.3d 144, 153 n.3 (2d Cir. 2007) ("[T]he [false advertising] plaintiff must also demonstrate that the false or misleading representation involved an inherent or material quality of the product.").

70. 15 U.S.C. § 1125(a)(1)(A) (2006).

71. 15 U.S.C. § 1125(a) (2006); *see, e.g.*, Mut. of Omaha Ins. Co. v. Novak, 836 F.2d 397, 400, 403 (8th Cir. 1987) (citing survey indicating that "ten percent of all the persons surveyed thought that Mutual 'goes along' with" defendant's use in upholding judgment that T-shirt with phrase "Mutant of Omaha" infringed "Mutual of Omaha" mark); Robert C. Denicola, *Freedom to Copy*, 108 YALE L.J. 1661, 1668 (1999) ("If trademark owners win enough high-profile cases or brag loudly enough about licensing revenues from ornamental use, consumers will naturally think that the products they see must be licensed, which in turn will help insure that a license is indeed required.").

72. McCarthy, *supra* note 3, at 46 (describing congressional amendments of section 43(a) as a "stamp of approval" on judicial interpretations).

73. Indeed, this vagueness effectively expands the power of trademark by arming markholders with plausible (or, at least, plausible-sounding) threats of suit in contexts far

B. Lagging Trademark "Defenses"

Before turning to the question of whether the modern Lanham Act accommodates new defenses as well as it does liability, we should first consider why new defensive doctrines may be needed.

A range of doctrines already limit trademark's scope. These safeguards protect both the competitors of trademark holders and the consuming public at large. In particular, they correct the tendency of trademark jurisprudence to neglect the interests of non-confused consumers. A successful trademark claim protects a channel of communication between the trademark holder and her customers, but it simultaneously closes a similar channel belonging to the junior user. The lost communications often have real value to non-confused consumers, so trademark liability may come at their expense. Trademark defenses and related doctrines mitigate these costs.[74]

Some defensive doctrines protect consumer access to information.[75] Others safeguard a competitive marketplace by preventing the lock-up of functional designs under the guise of protecting the source-identifying function of product design and trade dress.[76] Courts have also developed doctrines based on the First Amendment to prevent trademark law from threatening expressive rights (and, in turn, the right of listeners to hear the speaker's message).[77]

But while trademark's expanding scope is characterized by malleable

---

afield from traditional trademark infringement. The *in terrorem* prospect of a suit may suffice to deter legal trademark uses. *See, e.g.*, James Gibson, *Risk Aversion and Rights Accretion in Intellectual Property Law*, 116 YALE L.J. 882, 913 (2007).

74. For a fuller discussion of this aspect of doctrines that limit trademark's scope, see Grynberg, *supra* note 10, at 78-87.

75. So, for example, junior users are permitted to use trademarked words in their descriptive sense in the marketplace. *See* 15 U.S.C. § 1115(b)(4) (2006); *see generally infra* Section II.B.1 (discussing the classic fair use doctrine). Similarly, the first-sale, or exhaustion, doctrine protects the ability of the reseller of goods to do so notwithstanding the presence of affixed trademarks. *See, e.g.*, Davidoff & CIE, S.A. v. PLD Int'l Corp., 263 F.3d 1297, 1301 (11th Cir. 2001) ("The resale of genuine trademarked goods generally does not constitute infringement. . . . Under what has sometimes been called the 'first sale' or 'exhaustion' doctrine, the trademark protections of the Lanham act are exhausted after the trademark owner's first authorized sale . . . .").

76. *See* 15 U.S.C. § 1115(b)(8) (2006).

77. *See, e.g.*, Rogers v. Grimaldi, 875 F.2d 994, 999 (2d Cir. 1989) ("[T]he [Lanham] Act should be construed to apply to artistic works only where the public interest in avoiding consumer confusion outweighs the public interest in free expression."). In addressing a Lanham Act claim based on a movie title, *Rogers* adopted a balancing test that asks whether the title used by the defendant is artistically relevant to the underlying work and, if so, whether the use explicitly misleads as to source or content. *Id.*

standards of liability—i.e., a broad and vague conception of actionable confusion—traditional trademark defenses are comparatively rigid. Trademark law lacks any mechanism with the apparent flexibility of copyright's fair use defense.[78] The result may be an imbalance in trademark doctrine. Trademark liability expands to new realms without the counterbalance of translated defensive doctrines that are harder to apply in novel settings.[79]

For example, plaintiffs have used the Lanham Act to attack advertising that evokes the identity of unconsenting celebrities. In *Waits v. Frito-Lay*,[80] the Ninth Circuit allowed a false endorsement claim under section 43(a) by singer Tom Waits against the commercial use of a sound-alike musician.[81] The opinion's treatment of Waits's claim reflects the common

---

78. 17 U.S.C. § 107 (2006) (setting forth multifactor test to determine whether use of a copyrighted work is a protected fair use). Copyright's fair use doctrine was codified with the intent that courts would continue the common law development of the concept. Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 577 (1994) ("Congress meant § 107 'to restate the present judicial doctrine of fair use, not to change, narrow, or enlarge it in any way' and intended that courts continue the common-law tradition of fair use adjudication." (quoting H.R. REP. NO. 94-1476, at 66 (1976), *as reprinted in* 1976 U.S.C.C.A.N. 5659, 5679); S. REP. NO. 94-473, at 62 (1975)).

79. In one respect, trademark defenses are flexible in a manner detrimental to defendants. Many defensive doctrines incorporate the malleable likelihood of confusion standard into the determination of whether such confusion should be excused. While the Supreme Court has held that the classic fair use doctrine, codified in section 33(b)(4) of the Lanham Act, acts as a defense even where a likelihood of confusion exists, it held open the door for courts to consider the extent of any confusion in determining whether the defense applies. KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc., 543 U.S. 111, 123 (2004).

The expressive use test of *Rogers*, 875 F.2d 994, is another example. The Second Circuit announced the test as a way to mediate First Amendment and trademark policies when a trademark is used as the title of an expressive work. If the title is artistically relevant to the work, the test asks whether the defendant has done anything to mislead. This inquiry invites the return of confusion considerations. *See, e.g.*, Westchester Media v. PRL USA Holdings, Inc., 214 F.3d 658, 664-65 (5th Cir. 2000) (noting that under the Second Circuit applies the likelihood of confusion test to evaluate artistically relevant titles but that "the likelihood of confusion must be 'particularly compelling' to outweigh the First Amendment interests at stake") (citation omitted).

As Bill McGeveran notes, the degree to which likelihood of confusion considerations infect existing trademark defenses serves to strip potential defendants of any ex ante confidence regarding their litigation prospects. Even when defenses will ultimately protect socially beneficial uses, they may be unable to resolve cases at an early stage of the proceedings, therefore failing to provide prospective defendants with needed certainty. McGeveran, *supra* note 13, at 110-15.

80. 978 F.2d 1093 (9th Cir. 1992).

81. *Id.* The claim was made more viable by Waits's distinctive sound. He sounds like someone who "drank a quart of bourbon, smoked a pack of cigarettes and swallowed

law/congressional ratification dynamic discussed above. Even though Waits's singing style did not fit the traditional parameters of a trademark—and notwithstanding the questionable basis of false endorsement claims under the pre-amendment text of section 43(a)[82]—the court treated the defendant's practice as "the misuse of a trademark."[83] To get there, however, the panel did not analyze the statutory text in any great detail. It relied instead on opinions from outside the circuit that had permitted false endorsement claims under the Lanham Act.[84] These expansionist interpretations were supported by Congress's then-recent amendment of section 43(a), which the court recognized as codifying many of the "broad" judicial interpretations on which Waits's claim depended.[85]

But if a celebrity's identity is treated like a trademark,[86] then it stands to reason that traditional trademark defenses would apply to any infringement claims. Not so. In *Abdul-Jabbar v. General Motors Corp.*,[87] the retired basketball star sued the carmaker for a commercial that included the trivia question, "Who holds the record for being voted the most outstanding player of th[e NCAA men's basketball] tournament?" The answer: "Lew Alcindor, UCLA, '67, '68, '69."[88] This informative tidbit was followed by the question and answer: "Has any car made the 'Consumer Digest's Best Buy' list more than once?" "The Oldsmobile Eighty-Eight has."[89]

Abdul-Jabbar claimed the potential implication of endorsement vi-

---

a pack of razor blades . . . . Late at night. After not sleeping for three days." *Id.* at 1097.

82. These claims seemed to apply only to designations of geographic origin. *See supra* notes 40-43 and accompanying text.

83. *Waits*, 978 F.2d at 1110.

84. *Id.* at 1106-07.

85. *Id.* at 1107 ("[W]e read the amended language to codify case law interpreting section 43(a) to encompass false endorsement claims."). The court further contended that the legislative history of the 1988 amendments "makes clear that in retaining the statute's original terms 'symbol or device' in the definition of 'trademark,' Congress approved the broad judicial interpretation of these terms to include distinctive sounds and physical appearance." *Id.* (citations omitted).

86. Although the Ninth Circuit applies the multifactor likelihood of confusion test in such cases, it has "adapted these factors so as to be applicable to the celebrity cases." Downing v. Abercrombie & Fitch, 265 F.3d 994, 1007 (9th Cir. 2001).

87. 85 F.3d 407 (9th Cir. 1996).

88. *Id.* Abdul-Jabbar officially changed his name from Alcindor in 1971 and endorsed products under his new name. *Id.* at 409.

89. *Id.* During an ensuing film clip of the car, the ad further boasted "In fact, it's made that list three years in a row. And now you can get this Eighty-Eight special edition for just $18,995." The clip concluded with the printed messages, "A Definite First Round Pick," and "Demand Better, 88 by Oldsmobile" and voiceover, "It's your money." *Id.*

olated section 43(a).[90] His problem was that he had long since changed his name from Alcindor to Abdul-Jabbar. Insofar as he was making a trademark claim, he therefore appeared subject to the traditional trademark limitation that he not abandon his mark.[91]

Despite its comfort with employing the standard likelihood of confusion factors to establish the defendant's potential liability,[92] the court balked at importing an abandonment defense to an endorsement claim, concluding that to do so would "stretch" federal trademark law.[93] Trademark analogies could only be taken so far.[94]

My point is not that the abandonment defense should or should not have played a factor in *Abdul-Jabbar*.[95] Rather, unfair competition cases that accord trademark-like protections to nontraditional subject matter, like a celebrity's commercial identity, may not leave similar room for trademark-like defenses. Here, endorsement claims jumped the section 43(a) barrier, but the abandonment defense did not come along for the ride. The same dynamic appears in applications of defenses that are more

---

90. *Id.* at 410. The endorsement claim had stronger statutory support than that in *Waits* because it arose under the amended (and current) text of the Act. *See* 15 U.S.C. § 1125(a) (2006) (providing cause of action for use of any "device" that is likely to cause confusion or mistake as to "approval" of one's "goods, services, or commercial activities by another person").

91. A mark is abandoned when "its use has been discontinued with intent not to resume such use" or "any course of conduct of the owner, including acts of omission as well as commission, causes the mark to become the generic name for the goods or services on or in connection with which it is used or otherwise to lose its significance as a mark." 15 U.S.C. § 1127 (2006). This limitation is codified as a defense to a claim of infringement of a registered mark. 15 U.S.C. § 1115(b)(2) (2006). Non-use of a mark for three years amounts to prima facie evidence of abandonment. 15 U.S.C. § 1127. Abdul-Jabbar had not used the Alcindor name for a commercial purpose in over ten years. *Abdul-Jabbar*, 85 F.3d at 409.

92. *Id.* at 413.

93. *Id.* at 411 ("One's birth name is an integral part of one's identity; it is not bestowed for commercial purposes, nor is it 'kept alive' through commercial use.").

94. The court elaborated:

> In other words, an individual's given name, unlike a trademark, has a life and a significance quite apart from the commercial realm. Use or nonuse of the name for commercial purposes does not dispel that significance. An individual's decision to use a name other than the birth name-whether the decision rests on religious, marital, or other personal considerations-does not therefore imply intent to set aside the birth name, or the identity associated with that name.

Id. at 412.

95. Though one could argue on the underlying merits that identification of a celebrity by a name that he no longer wished to use in his personal life belies any false suggestion of endorsement.

central to consumer interests.

1. *"Classic" Fair Use*

Both the Lanham Act and the common law have long protected the right to use trademarked words in their descriptive sense.[96] Yet, the fair use doctrine's status as a true affirmative defense under the Lanham Act remained in doubt until recently.[97]

The defense is narrow. In its statutory incarnation, it extends only to use that "is descriptive of and used fairly and in good faith only to describe the goods or services of such party, or their geographic origin" and is a use "otherwise than as a mark."[98] Each element ("descriptive," "used fairly and in good faith," "only to describe," and "otherwise than as a mark") is essential to the defense. While the "descriptive" limitation is capable of an expansive, pro-defendant reading,[99] it is likewise capable of a narrow one, particularly when read in conjunction with the other elements, which are malleable enough to render the defense inapplicable if a court so wishes.[100]

Fair use's limitations stymied an effort to use the defense to check the expansion of the initial interest confusion doctrine to the Internet.[101] *Brookfield Communications v. West Coast Entertainment Corp.* involved the trademarked term "moviebuff," which described a searchable database

---

96. 15 U.S.C. § 1115(b)(4); *see also* 2 MCCARTHY, *supra* note 26, § 11:45 ("A junior user is always entitled to use a descriptive term in good faith in its primary, descriptive sense other than as a trademark.").

97. *KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc.* made clear that fair uses may be excused even if they cause some consumer confusion. 543 U.S. 111, 121-22 (2004). The Court, however, held open the possibility that the amount of confusion may have a bearing on whether the defense may be invoked. *Id.* at 123 ("It suffices to realize that our holding that fair use can occur along with some degree of confusion does not foreclose the relevance of the extent of any likely consumer confusion in assessing whether a defendant's use is objectively fair.").

98. 15 U.S.C. § 1115(b)(4). The statutory provision also defends the rights of those who share a name with a trademarked term to accurately describe themselves.

99. *See infra* note 106.

100. *See, e.g.*, *infra* note 123 and accompanying text; Grynberg, *supra* note 10, at 69. The malleability problem is further exacerbated in courts, like the Ninth Circuit, that allow the existence of some likelihood of confusion to influence its conclusion as to whether the defense applies in the first instance. *See* KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc., 408 F.3d 596, 609 (9th Cir. 2005).

101. "Initial interest confusion is customer confusion that creates initial interest in a competitor's product. Although dispelled before an actual sale occurs, initial interest confusion impermissibly capitalizes on the goodwill associated with a mark and is therefore actionable trademark infringement." Playboy Enters., Inc. v. Netscape Commc'ns Corp., 354 F.3d 1020, 1025 (9th Cir. 2004) (footnotes omitted).

of entertainment-related information.[102] The Ninth Circuit held that use by a competitor of the mark in the hidden text of its website created initial interest confusion because of the risk that web surfers might click on the "wrong" link when following search engine results.[103]

Whatever the harm to consumers of such diversions, they also carry potential benefits by enabling competitors to describe themselves as offering similar products or services to those of the trademark owner. So, for example, consumers who are unaware of the existence of generic acetaminophen may benefit if a search for "TYLENOL" returns the websites of drug sellers who offer a generic alternative and arrange to have their sites displayed in response to a search for the trademarked term.[104]

Though such uses honor the policies behind the classic fair use, they do not strictly comply with the letter of the doctrine. Accordingly, *Brookfield* rejected the defense with respect to the moviebuff mark. Because there was no space between the terms "movie" and "buff," the defendant was not using the term in its "pure" descriptive sense (e.g., "the website for movie buffs").[105]

To be sure, more flexible judicial interpretations of classic fair use exist in the case law.[106] But nothing in the text of section 33(b)(4) invites a flexible interpretation in the defendant's favor—rather, the multiple requirements of the defense suggest the contrary. More importantly, nothing in the statute gives prospective defendants any certainty about whether to proceed in the face of a trademark holder's threat of suit.[107]

---

102. Brookfield Commc'ns, Inc. v. W. Coast Entm't Corp., 174 F.3d 1036 (9th Cir. 1999).

103. *Id.* at 1062 (despite lack of source confusion "there is nevertheless initial interest confusion in the sense that, by using 'moviebuff.com' or 'MovieBuff' to divert people looking for 'MovieBuff' to its web site, West Coast improperly benefits from the goodwill that Brookfield developed in its mark").

104. *See generally* Grynberg, *supra* note 10, at 104-07 (discussing overlooked benefits of purported initial interest confusion). Nominative fair use may also address this problem in a limited way. *See infra* Section IV.C.

105. *Brookfield*, 174 F.3d at 1066 ("Even though ['MovieBuff'] differs from 'Movie Buff' by only a single space, that difference is pivotal. The term 'Movie Buff' is a descriptive term, which is routinely used in the English language to describe a movie devotee. 'MovieBuff' is not. . . .").

106. McGeveran, *supra* note 13, at 87-88 (discussing cases in which the defense is any use of trademarked term in its "descriptive sense" and not restricted to a use that describes defendant's product); *see, e.g.*, Packman v. Chicago Tribune Co., 267 F.3d 628 (7th Cir. 2001) (finding that a newspaper's sale of memorabilia displaying reproduced newspaper headline "Joy of Six" to celebrate championship of Chicago Bulls constituted fair use).

107. McGeveran, *supra* note 13, at 88 ("[I]nconsistent readings of § 33(b)(4) make it

2. *Genericism*

The problem raised by *Brookfield* also reflects the limitations of trademark's genericism doctrine, which serves a similar role to the classic fair use defense. Trademark law does not allow the protection of generic terms,[108] and trademarks that become generic lose protection.[109] The generic bar preserves the availability of certain market-useful terms (e.g., wine, book, car). However, the doctrine lacks a clear mechanism for protecting uses of another's mark in a generic way *even though* the mark retains distinctiveness. Returning to the example of a generic drug seller's use of TYLENOL as a keyword, such use may serve as the junior user's signal that "I am in the same product category as TYLENOL," not "I am TYLENOL." Permitting this kind of use would provide consumers with the same information benefits that genericism is intended to protect. Yet traditional generic mark doctrine, which pertains to the protectability of TYLENOL in the first instance, is not designed to accommodate this situation. Nor did *Brookfield* carve any new doctrinal space for such considerations.[110]

3. *Functionality and the Question of Aesthetics*

The bar to trademarking "functional" matter is another traditional marketplace protection that is now an enumerated defense to trademark infringement.[111] The functionality doctrine protects market competition by

---

difficult to tell in advance whether the defense will be available, especially if it is unclear in which circuit an eventual lawsuit might be brought.").

108. *See, e.g.*, Canal Co. v. Clark, 80 U.S. 311, 323 (1871) ("Nor can a generic name . . . be employed as a trade-mark and the exclusive use of it be entitled to legal protection.")

109. *See, e.g.*, 15 U.S.C. § 1064(3) (2006) (allowing mark cancellation petitions "if the registered mark becomes the generic name for the goods or services, or a portion thereof"); 15 U.S.C. § 1065(4) (2006) (denying incontestable status to marks that have become generic); 15 U.S.C. § 1127 (2006) (deeming abandoned marks when "any course of conduct of the owner . . . causes the mark to become the generic name for the goods or services").

110. *See* Grynberg, *supra* note 10, at 86. Indeed, *Brookfield*'s embrace of initial interest confusion undermines other defenses like the first-sale doctrine. *See* Std. Process, Inc. v. Total Health Disc., Inc., 559 F. Supp. 2d 932, 938-39 (E.D. Wis. 2008) (denying summary judgment to product reseller's first-sale and nominative use defenses in part on grounds that use of product's trademarked term in keyword advertising may cause consumer confusion).

111. 15 U.S.C. § 1115(b)(8) (2006); *see also* 15 U.S.C. § 1052(e)(5) (2006) (forbidding registration of trademark that "comprises any matter that, as a whole, is functional"); 15 U.S.C. § 1125(a)(3) (2006) ("In a civil action for trade dress infringement under this chapter for trade dress not registered on the principal register, the person who asserts trade dress protection has the burden of proving that the matter sought to be protected is

preventing trademark protection of useful product features even if such features perform a source-identifying function.[112]

Functionality has traditionally focused on excluding utilitarian features from trademark.[113] Although trademarking aesthetically pleasing features poses some of the same problems as protecting utilitarian designs, some judges fear the consequences of excluding aesthetic design from protection, lest sellers be deterred from bringing attractive goods to market.[114] Courts sometimes finesse the issue by redefining what might appear to be an aesthetic feature as a utilitarian one.[115]

When push comes to shove, the functionality defense often fails outside of its paradigmatic utilitarian case.[116] In *Au-Tomotive Gold, Inc. v.*

---

not functional."); *see generally* TrafFix Devices, Inc. v. Mktg. Displays, Inc., 532 U.S. 23 (2001) (setting forth tests for functional matter). The policy concerns animating the functionality bar are venerable. *See, e.g.*, Canal Co. v. Clark, 80 U.S. 311, 323 (1871) (observing that "[n]o one can claim protection for the exclusive use of a trade-mark or trade-name which would practically give him a monopoly in the sale of any goods other than those produced or made by himself," otherwise "the public would be injured rather than protected, for competition would be destroyed.").

112. *See, e.g.*, *TrafFix*, 532 U.S. at 35 (withholding trademark protection from a spring system designed to prevent roadside signs from blowing down in the wind).

113. *TrafFix* set forth two functionality tests. Under the so-called traditional test, " 'a product feature is functional,' and cannot serve as a trademark, 'if it is essential to the use or purpose of the article or if it affects the cost or quality of the article.' " *Id.* at 32 (quoting Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 165 (1995)). Under the second test, "a functional feature is one the 'exclusive use of which would put competitors at a significant non-reputation-related disadvantage.' " *Id.* (quoting *Qualitex*, 514 U.S. at 165). This latter test is often at play in aesthetic functionality situations.

114. *See, e.g.*, WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 199-200 (2003) (contending that aesthetic features used as marks do not disadvantage other firms when such features do not become "an attribute of the product" in a consumer's mind); Alex Kozinski, *Trademarks Unplugged*, 68 N.Y.U. L. REV. 960, 970 (1993) ("Allowing unrestricted copying of the Rolex trademark will make it less likely that Rolex, Guess, Pierre Cardin, and others will invest in image advertising, denying the image-conscious among us something we hold near and dear."). This concern begs the question of whether such incentives are a legitimate concern of trademark—as opposed to copyright or design patent—law.

115. Eco Mfg. LLC. v. Honeywell Int'l, Inc., 357 F.3d 649, 654 (7th Cir. 2003) (positing reasons why a circular thermostat shape may be functional independent of consumer aesthetic preference); *cf.* Brunswick Corp. v. British Seagull Ltd., 35 F.3d 1527, 1531 (Fed. Cir. 1994) (ruling that use of the color black for outboard motors was not entitled to trademark protection because "the color black exhibits both color compatibility with a wide variety of boat colors and ability to make objects appear smaller," creating a competitive need for the color by other engine manufacturers).

116. Despite the room in Supreme Court precedent for considerations of aesthetic functionality, *see supra* note 113, courts have been reluctant to embrace robust theories of aesthetic functionality. 1 MCCARTHY, *supra* note 26, § 7:80 (surveying federal circuit

Volkswagen of America, Inc.,[117] the holders of the Volkswagen and Audi trademarks sued the maker of automobile accessories who used the marks in key chains and license plate covers. The defendant argued that the logos were aesthetically functional aspects of the defendant's product.[118] That is, Volkswagen owners do not purchase a key with the VW logo because they think Volkswagen produced the key chain; they simply want their key chain to match their car.[119]

Note that the plaintiffs' cause of action depended on a prior expansion of trademark law beyond its traditional boundaries. No trademark claim would have been possible if courts had not expanded the likelihood of confusion concept to encompass situations in which the mark serves primarily not as a designator of source, but as the product itself (e.g., a Boston Celtics jersey, for which demand is for a Celtics jersey regardless of the physical source of the jersey).[120]

The Ninth Circuit had little difficulty embracing the expansion, concluding that the case presented an "easy analysis" with respect to likelihood of confusion.[121] Despite absence of any evidence of source confusion, the court weighed most of the multifactor test in favor of the plaintiffs[122] and found the defendant's business model—the legality of which the case was testing—to be in and of itself evidence of bad faith.[123] Similarly, the court relied on the theory of post-sale confusion of non-purchasers to dis-

---

courts and observing that most have either explicitly rejected aesthetic functionality or expressed doubts as to its validity).

117. Au-Tomotive Gold, Inc. v. Volkswagen of Am., Inc., 457 F.3d 1062 (9th Cir. 2006).

118. *Id.* at 1064.

119. *Id.*

120. *See supra* Section II.A.2.

121. *Au-Tomotive Gold*, 457 F.3d at 1076.

122. Some factors, like similarity of marks, *id.*, being especially easy once the underlying theory of litigation was accepted.

123. The court stated:

> Auto Gold knowingly and intentionally appropriated the exact trademarks of Volkswagen and Audi. Auto Gold argues, however, that it does not "intend" to deceive the public as to the source of the goods, but merely sought to fill a market demand for auto accessories bearing the marks. This argument is simply a recasting of aesthetic functionality. Even if we credit Auto Gold's proffered lack of intent, the direct counterfeiting undermines this argument. This factor tips against Auto Gold.

Id.; *cf. id.* at 1064 ("Auto Gold's incorporation of Volkswagen and Audi marks in its key chains and license plates appears to be nothing more than naked appropriation of the marks. The doctrine of aesthetic functionality does not provide a defense against actions to enforce the trademarks against such poaching.").

count the defendant's use of disclaimers.[124]

No similar flexibility accompanied the court's analysis and rejection of the aesthetic functionality defense. Part of the problem stemmed from the fact that in the typical case, the functionality defense is strong medicine. The defendant's contention normally is that the mark in question does not merit protection.[125] But the defendant could have made a different conceptual claim in *Au-Tomotive Gold*. That the carmaker logos were functional *in context* need not have undermined their distinctiveness as marks. For the panel, however, the broad strokes of the functionality doctrine suggested an extreme outcome—the loss of trademark protection for Volkswagen's and Audi's logos—which would have meant the "death knell for trademark protection."[126]

What is telling is that the court showed no inclination to adjust trademark defenses to accommodate a practice that arguably offers consumers the benefits of enhanced price competition. The court never considered the prospect that the functionality doctrine could be malleable enough to recognize functionality in the limited context of an adjacent market *without endangering* the protectability of the plaintiffs' marks in their core markets—where the marks *do* perform a source-identifying function—because to do so would render the "aesthetic function . . . indistinguishable from and tied to the mark's source-identifying nature."[127] While the court welcomed a trademark claim that reached well beyond confusion as to source,[128] the perceived threat to source identification of an adjusted functionality defense proved too much to contemplate.

---

124. *Id.* at 1077-78 ("Shorn of their disclaimer-covered packaging, Auto Gold's products display no indication visible to the general public that the items are not associated with Audi or Volkswagen. The disclaimers do nothing to dispel post-purchase confusion.").

125. *Cf.* 15 U.S.C. § 1125(a)(3) (2006) (placing burden on plaintiff in cause of action for infringement of unregistered trade dress to demonstrate that claimed trade dress is not functional).

126. *Au-Tomotive Gold*, 457 F.3d at 1064.

127. *Id.* at 1074; *see also id.* at 1073 ("[T]he rule [defendant] advocates injects unwarranted breadth into our caselaw. . . . In practice, aesthetic functionality has been limited to product features that serve an aesthetic purpose wholly independent of any source-identifying function.").

128. "A [l]ikelihood of confusion exists when customers viewing [a] mark would probably assume that the product or service it represents *is associated with the source* of a different product or service identified by a similar mark." *Id.* at 1075-76 (quoting Fuddruckers, Inc. v. Doc's B.R. Others, Inc., 826 F.2d 837, 845 (9th Cir. 1987) (alterations in original) (internal quotations omitted); *see also supra* notes 120-124 and accompanying text.

4. *Other Defenses*

The Lanham Act codifies a number of other trademark defenses. They exist to police the conduct of the trademark claimant and rarely implicate the issues discussed above.[129] Other "defenses" are either closely tied to the underlying cause of action, like nominative fair use, or were derived from extra-trademark sources, specifically the First Amendment. They present their own difficulties and are discussed in greater detail below.[130]

C. Summary

The modern equilibrium between the Lanham Act causes of action and trademark defenses disadvantages consumers. At its best, trademark law protects consumers and sellers alike. But overzealous mark protection may harm consumers by depriving them of valuable information or the benefits of market competition. Trademark defenses vindicate these interests, but rigid judicial interpretation often limits their effectiveness. For example, in *Au-Tomotive Gold*, the trademark holders were able to expand their monopoly to an adjacent market and deprive consumers of effective price competition because the asserted defense proved less flexible than the cause of action.[131] The court failed to analyze the interests of non-confused consumers in a meaningful way because there was no clear cut doctrinal box in which to place them, notwithstanding the existence of the functionality doctrine.

The Ninth Circuit's reluctance to flexibly interpret aesthetic functionality suggests that courts are hardly eager to create new trademark defenses.[132] To be sure, the current judicial attitude towards such creativity

---

129. *See generally* 15 U.S.C. § 1115(b)(4) (2006) (providing for defenses of fraudulent registration, abandonment, misrepresentation, prior use, violation of the antitrust laws, and indicating where equitable principles—such as laches, estoppel, and acquiescence—apply). *See also* notes 86-95 and accompanying text (discussing inapplicability of abandonment defense to *Abdul-Jabbar*).

130. *See infra* Part IV.

131. *See, e.g.*, Grynberg, *supra* note 10, at 85-86.

132. Indeed, reluctance to create new defenses may bias the evaluation of an infringement claim. *McNeil Nutritionals, LLC v. Heartland Sweeteners, LLC*, 511 F.3d 350 (3d Cir. 2007), involved a trade dress claim by the seller of the artificial sweetener Splenda (the brand name for sucralose) against the distributor of chemically equivalent store brands (e.g., sucralose sold under the Food Lion label in Food Lion stores). Although the trade dress of these brands shared similar coloring to Splenda, the presence of a distinct mark on the packaging of some of the store brands counteracted any likelihood of confusion. *Id.* at 360-61. Other packaging lacked such a prominent distinguishing feature, and the district court (applying the multifactor likelihood of confusion test) weighed the similarity of the marks in favor of the plaintiffs. *Id.* at 363 n.4. With respect to that packaging, the Third Circuit reversed the finding of no likelihood of confusion. *Id.* at 367

says little about judicial freedom to act should such attitudes evolve. The next Part turns to the question of whether such leeway exists.

III. A WAY OUT? TRADEMARK'S COMMON LAW PROBLEM

Arguments that trademark's scope is too broad are nothing new,[133] as are proposals for new or revitalized doctrines to check it. In recent years, courts[134] and commentators[135] alike have considered the existence and reach of a trademark "use" requirement and debated the extent to which the Lanham Act[136] or older common law sources[137] even contain the requirement. What is striking is the shared confidence of the debaters that courts have the power to craft new defenses to liability.[138] Their faith jibes well with trademark's history. Trademark liability expanded in the first instance through judges acting in a "common law" manner, even after enactment of a comprehensive federal statute in 1946. Given trademark's common law roots, why wouldn't that tradition enable and guide the development of defenses to complement and check any overgrowth of lia-

("[T]here is no way the District Court could have ultimately balanced the *Lapp* factors against McNeil after weighing the first, second, seventh, eighth, and ninth *Lapp* factors in its favor."). Of interest is the court's preclusion of any reconsideration on remand because it feared that the lower court was attempting to create a new defense to infringement. In ruling against the plaintiff, the district court noted that consumer awareness of the existence of store brand products would negate any likely confusion especially when coupled with other signals like price differential and shelf location. *Id.* For the court of appeals this observation treaded too close to a categorical defense and thus justified taking the matter out of the trial court's hands.

> The danger in the District Court's result is that producers of store-brand products will be held to a lower standard of infringing behavior, that is, they effectively would acquire *per se* immunity as long as the store brand's name or logo appears somewhere on the allegedly infringing package, even when the name or logo is tiny. The Lanham Act does not support such a *per se* rule.

Id. at 367-68.

133. *See, e.g.*, *supra* notes 5-8.

134. *Compare* Rescuecom Corp. v. Google Inc., No. 06-4881-cv, 2009 WL 875447 (2d. Cir. Apr. 3, 2009) (holding Google's sale of trademarked terms as keywords to be a use in commerce under the Lanham Act), *with* 1-800 Contacts, Inc. v. WhenU.com, Inc., 414 F.3d 400, 403 (2d Cir. 2005) (holding internet advertising service that supplied competitor pop-up ads when plaintiff's website was visited did not "use" plaintiff's trademarks).

135. *See supra* notes 8 and 14.

136. *See, e.g.*, Dogan & Lemley, *supra* note 8, at 1675-77.

137. *Id.* at 1677. *But compare, e.g.*, Dinwoodie & Janis, *Contextualism*, *supra* note 14, at 1609-22.

138. *See supra* notes 14-15 and accompanying text.

bility?

One answer is that times have changed. Whatever the past willingness of courts to treat statutes in an open-ended manner, ours is said to be a more self-consciously formalist age. Today commentators routinely note the prominence, if not triumph, of textualist interpretive theories that leave comparatively little room for judges to blunt harsh statutory edges with invocations of congressional intent, pragmatic considerations, or the discovery of statutory gaps in need of filling by federal common law.[139] Legislative history indicating Congress's intent for courts to "continue to interpret" section 43(a) may therefore not matter.[140]

Recent Supreme Court trademark jurisprudence is consistent with this storyline. Almost twenty years ago, *Two Pesos, Inc. v. Taco Cabana, Inc.* recognized expansive trademark rights in unregistered trade dress despite their questionable statutory pedigree.[141] While *Two Pesos* ratified earlier expansive judicial interpretation of the Lanham Act's scope, the Supreme Court acted against a background of congressional acquiescence and endorsement. Since then, the Court has used text-bound interpretations of the Lanham Act to slow further expansion of trademark's domain. The resulting opinions have checked further expansion of trademark's scope. But trademark had already come far, with many gains reinforced by open-ended statutory language. Because similar language does not exist with respect to trademark defenses, any "formalist shift" in trademark jurisprudence presents serious challenges for those who would rely on new defensive doctrines to curtail trademark's scope.

A. A Formalist Age?

Federal trademark law has always had a strong common law component. The pre-Lanham Act federal statute was limited, so much of the federal judicial action consisted of administering the pre-*Erie* federal common law of unfair competition.[142] This tradition continued after passage of the Lanham Act, and Congress largely endorsed the resulting de facto

---

139. *See infra* Section III.A.

140. S. REP. NO. 100-515 (1988), *as reprinted in* 1988 U.S.C.C.A.N. 5577, 5603. The quoted passage, moreover, does not give an indication that Congress intended to invite the creation of new defenses as opposed to the continued application of the trademark cause of action to novel settings, given that the report focused on the statute's role as filling "an important gap in federal unfair competition law." *Id.*

141. Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763 (1992); *see supra* notes 37-41 and accompanying text.

142. S. REP. NO. 79-1333 (1946), *as reprinted in* 1946 U.S.C.C.A.N. 1274, 1276-77 (citing the post-*Erie* lack of federal common law as reason for national trademark legislation); McCarthy, *supra* note 3, at 46-48.

"common law" of federal trademark protection by enacting more expansive statutory text.[143]

Had Congress remained silent, trademark's gains might be less secure. Numerous commentators argue that the current legal landscape is increasingly formalist.[144] That could mean any number of things,[145] but two aspects are of particular importance with respect to designing new trademark defenses. First, the increasing prominence of textualist statutory interpretation means that construction of statutory text is less likely than in the past to be guided by legislative history or pragmatic considerations.[146] Modern textualists emphasize, however, that statutory construction is not strictly limited to the text (and interpretive sources like contemporary dictiona-

---

143. *See supra* Section II.A.

144. *See, e.g.*, William N. Eskridge, Jr., *The New Textualism*, 37 UCLA L. REV. 621, 624 (1990) ("The new textualism is the most interesting development in the Court's legisprudence (the jurisprudence of legislation) in the 1980s."); Jerry L. Mashaw, *Textualism, Constitutionalism, and the Interpretation of Federal Statutes*, 32 WM. & MARY L. REV. 827, 832 (1991) ("Is textualism dominant, or at least a major new direction in the approach to statutory interpretation? My tentative answer is 'yes.' "); Jonathan T. Molot, *The Rise and Fall of Textualism*, 106 COLUM. L. REV. 1, 2 (2006) ("Textualists have been so successful discrediting strong purposivism, and distinguishing their new brand of 'modern textualism' from the older, more extreme 'plain meaning' school, that they no longer can identify, let alone conquer, any remaining territory between textualism's adherents and nonadherents."); Siegel, *supra* note 2, at 1057 ("[E]veryone must acknowledge the valuable and very significant achievement of Justice Scalia in recalling the attention of the legal community to the importance of text in statutory interpretation. In a significant sense, we are all textualists now."); Thomas C. Grey, *The New Formalism* (Stanford Pub. Law & Legal Theory Working Paper Series, Paper No. 4, 1999), *available at* http://ssrn.com/abstract=200732 ("It has long been an insult in sophisticated legal circles to call someone a formalist. . . . But within the last decade or so (overnight in jurisprudential time) this has changed . . . .").

145. For example, Grey identifies four formalist jurisprudential urges: objectivism (the desire for determinate rules); originalism in constitutional law; textualism (as opposed to statutory interpretation based on legislative purposes); and conceptualism, the desire for bodies of law like contract or tort to be treated "as coherent structures of concepts or principles." Grey, *supra* note 144, at 2.

146. John F. Manning, *What Divides Textualists from Purposivists?*, 106 COLUM. L. REV. 70, 110 (2006) [hereinafter, Manning, *What Divides?*] ("Properly understood, textualism means that in resolving ambiguity, interpreters should give precedence to semantic context (evidence about the way reasonable people use words) rather than policy context (evidence about the way reasonable people would solve problems)."). *Compare* Molot, *supra* note 144, at 23 ("In the immediate aftermath of the New Deal and legal realism, the Court's strong purposivism was perceived to be entirely compatible with legislative supremacy."). It bears noting that definitions of textualism are not free from debate. *See* Caleb Nelson, *What is Textualism?*, 91 VA. L. REV. 347, 351 (2005) ("[S]omeone seeking to predict how textualist judges will diverge from intentionalist judges is well-advised to start with the distinction between rules and standards.").

ries), but rather considers broader statutory context as well.[147] This caveat dovetails with the second formalist strand relevant to trademark defenses—the effort to harmonize discrete bodies of law into internally consistent wholes.[148]

Textualism's impact extends beyond the academy.[149] While Professor Calabresi, writing in 1982, could advocate judicial updating of statutes as a form of common lawmaking that had long been practiced without open acknowledgment,[150] Judge Calabresi, writing in 2006, agreed that his fa-

---

147. *See, e.g.*, Scalia, *supra* note 2, at 17 (emphasizing interpretation of " 'objectified' intent—the intent that a reasonable person would gather from the text of the law, placed alongside the remainder of the *corpus juris*"); Manning, *What Divides?*, *supra* note 146, at 79 ("In contrast with their ancestors in the 'plain meaning' school . . . modern textualists do not believe that it is possible to infer meaning from 'within the four corners' of a statute. Rather, they assert that language is intelligible only by virtue of a community's shared conventions for understanding words in context.") (footnote omitted).

148. Grey, *supra* note 144, at 2 (defining the "conceptualism" formalist tendency by explaining that formalists "prefer to treat abstract categories like contract and tort as coherent structures of concepts and principles, rather than as bodies of sublegislation generated in the course of judicial dispute-resolution"). While Grey defines this tendency with respect to common law categories (e.g., contract or tort), he makes clear that the impulse applies to bodies of law that have been reduced in whole or in part to statutory law. *Id.* at 24-25. He similarly fits Justice Scalia's jurisprudence into this model, claiming that for Scalia, fidelity to the derived overarching rules may even trump apparent statutory text to the contrary. *Id.* at 25 (discussing Antonin Scalia, *Assorted Canards of Contemporary Legal Analysis*, 40 CASE W. RES. L. REV. 581 (1990) [hereinafter, Scalia, *Assorted Canards*]). For Scalia, and the new formalists more generally, ad hoc judgments must do more than simply avoid contradiction. "The system must be intelligible and transparent as well as consistent. Consistency can check judges only to the degree that inconsistency can readily be identified. This will not be the case with a 'system' made up of thousands of independent ad hoc totality-of-the-circumstances determinations." *Id.* (footnotes omitted). For his part, Scalia argues:

> Without such a system of binding abstractions, it would be extraordinarily difficult for even a single judicial law-giver to be confident of consistency in his many *ad hoc* judgments; and it would be utterly impossible to operate a hierarchical judicial system, in which many individual judges are supposed to produce 'equal' protection of the laws.

Scalia, *Assorted Canards*, *supra*, at 589.

149. *See* Manning, *What Divides?*, *supra* note 146, at 109 n.141 (collecting examples); Molot, *supra* note 144, at 32-33 (arguing that "the broad appeal of textualism's underlying premises has led judges who do not consider themselves adherents to heed textualism's warnings about the pitfalls of strong purposivism and to alter their approach to statutory interpretation" and collecting empirical studies confirming effect) (footnotes omitted); *infra* note 152.

150. GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES 178-181 (1982); *id.* at 164 ("What, then, is the common law function to be exercised by courts today? *It is no more and no less than the critical task of deciding when a retentionist or a*

vored approach "is simply not a part of our legal system."[151] Consistent with these claims, commentators have observed that recent Supreme Court jurisprudence is less receptive to arguments based on legislative history or statutory purpose,[152] hostile to the use of the "federal common law,"[153]

---

revisionist bias is appropriately applied to an existing statutory or common law rule."); *id.* at 166 (arguing that adopting advocated approach "will only be recognizing the changes, not making them" and that this would merely be a "seeing of the world as it is").

151. Hayden v. Pataki, 449 F.3d 305, 367 (2d Cir. 2006) (Calabresi, J., dissenting). He explained:

> [S]ome scholars, myself included, have suggested that it might be a good idea if, as a starting point, in certain circumstances, courts were permitted to read the law according to what they perceived to be the will of the current Congress, rather than that of a long-gone-by one. But whatever the merits of such an arrangement in the abstract, it is simply not a part of our legal system.

Id. (citations omitted).

152. Thomas W. Merrill, *Textualism and the Future of the Chevron Doctrine*, 72 WASH. U. L.Q. 351, 355 (1994) (documenting Court's increasing use of dictionaries and decreasing use of legislative history). Writing in 1990, Eskridge observed:

> The Supreme Court has not thrown over its traditional approach to legislative history in favor of the new textualism, yet. In each year that Justice Scalia has sat on the Court, however, his theory has exerted greater influence on the Court's practice. This influence has been manifest in three respects. First, the Court is now somewhat less willing to refer to legislative history when the statutory text has a plain meaning. Second, the Court more often determines that a statutory text has a plain meaning by reference to structural textual arguments. Third, the Court has been increasingly influenced by textual and procedural canons of statutory interpretation.

Eskridge, *supra* note 144, at 656. Other studies have found similar trends, though some observe a comparatively small countermovement on legislative history led by Justices Breyer and Stevens. Molot, *supra* note 144, at 32 n.135 (collecting sources); *see also* Michael H. Koby, *The Supreme Court's Declining Reliance on Legislative History: The Impact of Justice Scalia's Critique*, 36 HARV. J. ON LEGIS. 369, 395 (1999) ("[T]here has emerged a clear and unmistakable pattern of decline in the use of legislative history by the Supreme Court. While the pattern is most acute in the decisions of more conservative justices, moderate and liberal justices are also citing to legislative history less often."). Other scholars challenge the categories used in such analyses as being incomplete. Jane S. Schacter, *The Confounding Common Law Originalism in Recent Supreme Court Statutory Interpretation: Implications for the Legislative History Debate and Beyond*, 51 STAN. L. REV. 1, 5 (1998) ("My analysis of the recent opinions suggests that these categories are far too stylized to capture the Court's interpretive practices which, in fact, cut across these familiar categories."); *see generally* FRANK B. CROSS, THE THEORY AND PRACTICE OF STATUTORY INTERPRETATION 134-39 (2008) (discussing research of the Court's methodology).

153. Henry Paul Monaghan, *Supreme Court Review of State-Court Determinations of State Law in Constitutional Cases*, 103 COLUM. L. REV. 1919, 1984 n.317 (2003) ("The Court has increasingly treated federal common law as a suspect enterprise, except within

and disinclined to read causes of action or defenses into statutes that do not clearly provide for them.[154] While most scholarship focuses on the Supreme Court, at least some evidence suggests that the textualist trend is equally, if not more, pronounced in the federal circuit courts.[155]

Ali v. Federal Bureau of Prisons, decided last year, offers a nice example of judicial practice following academic commentary. *Ali* addresses a bar to suit based on detention of property by "any officer of customs or excise or any other law enforcement officer."[156] The interpretive question is whether the phrase "any other law enforcement officer" literally means *any* law enforcement officer, or does the text limit the phrase to include only other officers when they are enforcing customs or excise laws? By a 5-4 vote, the Court took the first approach.

Of interest here is the relatively narrow battleground for the majority opinion and primary dissent. Their arguments were almost entirely textual, focusing on statutory context and application of interpretive canons like

---

the narrowest range."). Monaghan argues that this hesitation applies to use of federal common law to fill interstitial gaps in legislation. *See id.*

154. Last year, the Court declared it "settled that there is an implied cause of action only if the underlying statute can be interpreted to disclose the intent to create one." Stoneridge Inv. Partners, L.L.C. v. Scientific-Atlanta, Inc., 128 S. Ct. 761, 772 (2008); *see also* Burlington Indus., Inc. v. Ellerth, 524 U.S. 742, 755 (1998) (creating affirmative defense to liability under Title VII, but maintaining that "[t]he resulting federal rule . . . is statutory interpretation pursuant to congressional direction. This is not federal common law in the strictest sense.") (internal quotation omitted); Monaghan, *supra* note 153, at 1984 n.317 (citing *Burlington Indus.* in support of proposition that "the Court has gone to rather startling lengths to cast its results as statutory interpretation rather than federal common law"). Regardless of the majority's level of candor in *Burlington Indus.*, what is telling is its implicit agreement with the dissent that the creation of a defense based on policy alone would be illegitimate. *Cf. Burlington Indus.*, 524 U.S. at 772 (Thomas, J., dissenting) (characterizing the majority holding as "a product of willful policymaking, pure and simple").

155. Cross's own study of the Court's statutory interpretation cases between 1994 and 2002 reports that the Court uses textualism more than legislative intent, but by not as large a margin as the commentary would suggest, CROSS, *supra* note 152, at 145, and that all Justices demonstrated pluralist tendencies. *Id.* at 158. Of greater interest, perhaps, is Cross's analysis of interpretation practices in the circuit courts, which suggests that the " 'death of legislative history' as an interpretive tool is much more profound in the circuit courts than in the Supreme Court." *Id.* at 185; *id.* at 187 (speculating that the shift may be due to the "conventional wisdom" regarding the rise of textualism). At the same time, reliance on textualism increased. *Id.* at 188. Cross's analysis also indicates an increase in use of references to pragmatism, though he cautions that his tools for identifying such cases are relatively crude. *Id.* at 189 ("Pragmatism may be becoming more acceptable as an interpretive tool for the judiciary, though its absolute frequency is uncertain.").

156. Ali v. Fed. Bureau of Prisons, 128 S. Ct. 831, 834 (2008) (analyzing 28 U.S.C. § 2680(c)).

ejusdem generis and *noscitur a sociis*.[157] By contrast, Justice Breyer's separate dissent, which argued that the relevant context "extends well beyond Latin canons and other such purely textual devices" to invoke the statute's legislative history and pragmatic considerations, drew only the vote of Justice Stevens.[158]

Ali is obviously just one case, and the generalizations of the commentary are contestable.[159] This Article has no ambition of engaging, much

---

157. *Compare id.* at 838-41 (rejecting arguments based on *ejusdem generis*, *noscitur a sociis*, and the presumption against superfluity), *with id.* at 842-44 (Kennedy, J., dissenting) (arguing to the contrary). *Ejusdem generis* ("of the same kind or class") is the canon of construction providing that where a general term follows a group of specific ones, the general term should be interpreted to include only matters of the same type as encompassed by the prior specific terms. BLACK'S LAW DICTIONARY 556 (8th ed. 2004). *Noscitur a sociis* ("it is known by its associates") is the principle that ambiguous words draw meaning from surrounding words. *Id.* at 1087.

158. *Ali*, 128 S. Ct. at 849 (Breyer, J., dissenting); *see also id.* at 850-51 (making arguments based on legislative history and practical implementation). Even when Justice Breyer's relatively less textual approach carried the day in *Zuni Pub. Sch. Dist. No. 89 v. Dep't of Educ.*, 550 U.S. 81 (2007), his colleagues noted discomfort with his method. *Zuni* was a 5-4 decision turning on interpretation of debatably ambiguous language in the federal Impact Aid Act (a school financing statute), Justice Breyer began his main analysis by "depart[ing] from a normal order of discussion" by "first examin[ing] the provision's background and basic purposes" rather than its text. *Id.* at 90. While holding his majority, Breyer's departure provoked ridicule by the dissent:

> The opinion purports to place a premium on the plain text of the Impact Aid statute, but it first takes us instead on a roundabout tour of "[c]onsiderations *other* than language,"—page after page of unenacted congressional intent and judicially perceived statutory purpose. . . . This is a most suspicious order of proceeding. . . .

Id. at 108-09 (Scalia, J., dissenting) (emphasis added by Justice Scalia) (citations omitted). More tellingly, as a reflection of the Court's currently favored practices, Breyer's structure provoked a separate concurrence. Joined by Justice Alito, Justice Kennedy observed:

> In this case, the Court is correct to find that the plain language of the statute is ambiguous. It is proper, therefore, to invoke *Chevron*'s rule of deference. The opinion of the Court, however, inverts *Chevron*'s logical progression. Were the inversion to become systemic, it would create the impression that agency policy concerns, rather than the traditional tools of statutory construction, are shaping the judicial interpretation of statutes. It is our obligation to set a good example; and so, in my view, it would have been preferable, and more faithful to *Chevron*, to arrange the opinion differently. Still, we must give deference to the author of an opinion in matters of exposition; and because the point does not affect the outcome, I join the Court's opinion.

Id. at 107.

159. *See supra* note 152. Committed textualists have little difficulty identifying exceptions to the Supreme Court's textualist tilt. John F. Manning, *Justice Scalia and the*

less resolving, the attending debates.[160] And indeed, shifts in emphasis in interpretation methodology do not mean that judicial practices are not, on the whole, eclectic.[161] The claim here is not that the Court is or is not more formalist now than in the past. Nor is it necessary to argue that the Court is more formalist with respect to trademark litigation than in the past. My more modest contention is that since *Two Pesos*, the last time the Court construed the pre-1988 Lanham Act, all of the Court's important interpretations of the post-amendment statute have been consistent with the "formalist narrative" (specifically, its textual focus, and a desire to treat bodies of law as a unified whole) and incompatible with treating the Lanham Act as authorizing a federal common law of unfair competition. This is even true of those trademark opinions that appear to depart from the formalist line. These purported deviations are important, however, because they

---

Legislative Process, 62 N.Y.U. ANN. SURV. AM. L. 33, 42 n.35 (2006). Likewise, several articles documenting the rise of formalist tendencies in the courts are often at pains to suggest that the changes seen may not be so profound. *See, e.g.*, Grey, *supra* note 144, at 29 ("[A]t its theoretical core, the new formalism is just the old legal pragmatism, now mostly in the hands of conservatives rather than Progressives, New Dealers, and post-New-Deal liberals."); Siegel, *supra* note 2, at 1057 ("Although the battle over statutory interpretation has been waged with harsh words, the positions of the warring camps are not nearly as far apart as they might seem. Each of the competing methods of statutory interpretation accepts some of the insights of the others.") (footnote omitted). This is especially the case when combined with the observation that the newer forms of textualism are more open to contextual considerations than the old.

> In scholarship and case law alike, what one finds is convergence of opinion. On one hand, the purposivism that prevailed in prior decades has largely disappeared and textualist rhetoric has made its way into mainstream judicial opinions. On the other hand, even the most committed textualists have openly acknowledged that text can be ambiguous, that judges must read statutes in context, and that statutory purposes merit consideration in at least some cases. Ironically, at a time when the textualism debate seems to be garnering more attention—and even making its way into the mainstream press—that which unites textualists and purposivists seems to outweigh that which divides them.

Molot, *supra* note 144, at 35-36 (footnotes omitted). While maintaining that their points of emphasis still differ from non-textualists, Manning, *What Divides?*, *supra* note 146, at 76 (arguing that purposivists, unlike textualists, will allow "sufficiently pressing policy cues to overcome" semantic evidence), some self-described textualists partially concede these points, *id.* at 75-76 (agreeing that interpretation requires looking beyond text and statutory purpose is relevant to construction if not derived from legislative history).

160. I am not arguing that textualism is superior to or more legitimate than intentionalist or pragmatic schools of judging. Nor am I defending the judicial proponents of textualist methodology from claims that they apply their doctrine inconsistently or in a manner that favors a political agenda.

161. CROSS, *supra* note 152, at 158 (observing that while justices vary in emphasis of methodology, "most justices show pluralist tendencies").

show potential avenues for doctrinal innovation notwithstanding the Court's trademark formalism.

B. "Trademark Formalism" at the Supreme Court

Whatever can be said about the generalizations in the last section with respect to the judiciary as a whole, they jibe with the Supreme Court's trademark jurisprudence since *Two Pesos*. After embracing trademark's post-Lanham-Act-enactment expansion, the Court has since taken a skeptical view of further trademark expansion, using textualist rationales to guide their opposition. These opinions recognize that the Act—even the expansively worded section 43(a)—"does not have boundless application as a remedy for unfair trade practices."[162] Accordingly, the Court confined the dilution cause of action,[163] limited the reach of "reverse passing off" claims,[164] clarified that the classic fair use defense is a true affirmative defense,[165] broadened the functionality doctrine,[166] and tightened the requirements for obtaining protection of trade dress.[167] Contrary to *Two Pesos*, the Court refused to ratify expansionist rulings by other courts absent any statutory signals of congressional agreement. Instead, the built-in protections of the Lanham Act found a ready audience in a Court disposed to focus on textual considerations in deciding trademark cases.[168]

1. Textual Checks to Further Trademark Expansion

Trademark expansionism hit its Supreme Court peak in *Two Pesos*. Since then, further growth has received scant high-Court support with one exception. *Qualitex Co. v. Jacobson Products Co., Inc.* approved the use

---

162. Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23, 29 (2003) (quoting Alfred Dunhill, Ltd. v. Interstate Cigar Co., 499 F.2d 232, 237 (2d Cir. 1974)).

163. Moseley v. V Secret Catalogue, Inc., 537 U.S. 418, 433-34 (2003) (holding that the federal dilution statute required proof of actual dilution and not merely likelihood of dilution).

164. *Dastar*, 539 U.S. at 32-37 (2003) (holding the Lanham Act's prohibition of false designations of origin do not prohibit uncredited copying of another's work).

165. KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc., 543 U.S. 111, 121-22 (2004).

166. TrafFix Devices, Inc. v. Mktg. Displays, Inc., 532 U.S. 23, 32-35 (2001).

167. Wal-Mart Stores, Inc. v. Samara Bros., 529 U.S. 205, 211 (2000).

168. None of the cases discussed below drew a dissent, which is arguably consistent with a formalist narrative. RICHARD A. POSNER, HOW JUDGES THINK 50 (2008) (citing statistics indicating rising percentage of unanimous opinions in Supreme Court as possible evidence of increasing legalism among Justices, but proffering caveats and alternative explanations). To be sure, characterizing the Court's trademark jurisprudence based on the opinions of the Justices runs afoul of Judge Posner's contention that the formalist trappings of judicial opinions are more the product of law clerk drafting rather a true reflection of how the opinions are decided. *Id.* at 219-21.

of color as a trademark.[169] But this expansion, if it can be called an expansion,[170] found direct support in the Lanham Act's text, which provides that a trademark may be any "any word, name, symbol, or device, or any combination thereof."[171] Branding one's goods with color for source identification is at least arguably the use of a device.[172] Other efforts to expand trademark's scope had weaker foundations in the statutory text or structure, and they failed as a result.[173]

---

169. 514 U.S. 159 (1995).

170. The Court noted that its confidence in color as a mark only extended to situations in which the color had achieved secondary meaning (in much the same circumstances as when a descriptive mark may be a trademark). *Id.* at 163. Later precedent characterized *Qualitex* as holding that color requires secondary meaning to serve as a trademark. *Wal-Mart*, 529 U.S. at 212.

171. 15 U.S.C. § 1127 (2006).

172. Justice Breyer, true to the form discussed above, invoked background principles of trademark, but did so in conjunction with the statutory text:

> Both the language of the Act and the basic underlying principles of trademark law would seem to include color within the universe of things that can qualify as a trademark. The language of the Lanham Act describes that universe in the broadest of terms. It says that trademarks "includ[e] any word, name, symbol, or device, or any combination thereof." § 1127. Since human beings might use as a "symbol" or "device" almost anything at all that is capable of carrying meaning, this language, read literally, is not restrictive. The courts and the Patent and Trademark Office have authorized for use as a mark a particular shape (of a Coca-Cola bottle), a particular sound (of NBC's three chimes), and even a particular scent (of plumeria blossoms on sewing thread). *See, e.g.*, Registration No. 696,147 (Apr. 12, 1960); Registration Nos. 523,616 (Apr. 4, 1950) and 916,522 (July 13, 1971); *In re* Clarke, 17 U.S.P.Q. 2d 1238, 1240 (TTAB 1990). If a shape, a sound, and a fragrance can act as symbols why, one might ask, can a color not do the same?
>
> A color is also capable of satisfying the more important part of the statutory definition of a trademark, which requires that a person "us[e]" or "inten[d] to use" the mark "to identify and distinguish his or her goods, including a unique product, from those manufactured or sold by others and to indicate the source of the goods, even if that source is unknown." 15 U.S.C. § 1127.

Qualitex, 514 U.S. at 162.

173. *See also* Graeme B. Dinwoodie, *The Trademark Jurisprudence of the Rehnquist Court*, 8 MARQ. INTELL. PROP. L. REV. 187, 202 (2004) ("*Two Pesos* might be read as the high point of trade dress protection under this Court, with *Qualitex* hinting at both expansion and caution. In [later cases], the Court signaled a desire to rein in claims under the Lanham Act."). Professor Dinwoodie acknowledges the textualist aspects of the Court's change of heart, but is more skeptical than I am with respect to their predominance. He argues that "the Court's inconsistent use of textual interpretation (most notably between *Two Pesos* and *Wal-Mart*) shows that other considerations do inform the Court's analy-

a) Restricting Dilution

Moseley v. V Secret Catalogue, Inc.[174] unanimously rejected the claim that the Federal Trademark Dilution Act ("FTDA") requires only likelihood of dilution (in much the same manner that the Lanham Act's infringement cause of action only requires a likelihood of confusion). The reason was simple: the statute's terms at the time only explicitly reached acts that "cause[] dilution" and not those causing "likelihood" of dilution.[175] This fact sufficed to cabin the dilution cause of action.[176]

Perhaps the most jurisprudentially interesting aspect of *Moseley* is that the issue generated a circuit split in the first place. Although some circuits relied on statutory language to embrace an actual dilution standard,[177] others looked elsewhere. The Sixth Circuit's *Moseley* opinion, ultimately reversed by the Supreme Court, evaded the FTDA's text by interpreting the statute's legislative history to conclude that Congress intended a broad remedy, and was unlikely to have undermined that intention with the difficult standard of proof of an actual dilution requirement.[178] Similarly, the

---

sis." *Id.* at 207; *see also* Dinwoodie, *Defining Defenses*, *supra* note 13, at 143 n.182 (responding to draft version of this Article by arguing that the "formalist tendencies of the Supreme Court in recent cases are sufficiently coupled with more functionalist concerns not to discourage" efforts to create defenses). As I argue in this Section, however, the fact that non-textual factors may influence the Court when it chooses from among textually acceptable outcomes is a far cry from an invitation to render opinions in the common law style absent some reasonable textual basis.

174. 537 U.S. 418 (2003).

175. *Id.* at 433 ("This text unambiguously requires a showing of actual dilution, rather than a likelihood of dilution."); *see also id.* (looking to statutory definition of dilution to further support conclusion). Justice Stevens's opinion contained a brief section (not joined by Justice Scalia) on the dilution provision's legislative history, *id.* at 430-31, but the discussion had no bearing on the holding.

176. Until Congress could weigh in. Congress undid the Court's handiwork by passing the Trademark Dilution Revision Act of 2006, Pub. L. No. 109-312, 120 Stat. 1730 (codified as amended in scattered sections of 15 U.S.C.) (amending 15 U.S.C. § 1125(c) to make actionable acts causing a likelihood of dilution).

177. *See, e.g.*, Westchester Media v. PRL USA Holdings, Inc., 214 F.3d 658, 670 (5th Cir. 2000) ("[W]e endorse the Fourth Circuit's holding that the FTDA requires proof of actual harm since this standard best accords with the plain meaning of the statute.").

178. *V Secret Catalogue, Inc. v. Moseley*, 259 F.3d 464, 476 (6th Cir. 2001). It is worth noting, moreover, how generally worded that history was, considering the mileage the panel got out of it:

> As the Congressional Record indicates, dilution is "an injury that differs materially from that arising out of the orthodox confusion. Even in the absence of confusion, the potency of a mark may be debilitated by another's use. This is the essence of dilution. Confusion leads to immediate injury, while dilution is an infection, which if allowed to spread,

Second Circuit had recognized that the FTDA's language supported an actual dilution standard "in that it uses the formulation, 'causes dilution,' rather than referring to 'likelihood of dilution.' "[179] Nonetheless that court rejected an actual dilution requirement as "excessive literalism" where such a standard would permit injury without compensation.[180] At the end of the day, of course, literalism prevailed at the Supreme Court without dissent.

b) A Defense Is a Defense Is a Defense

KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc. held that the classic fair use defense, codified by section 33(b) of the Lanham Act, once established, excuses any liability for causing a likelihood of confusion.[181] Here again, the Lanham Act's text sufficed to rule that the fair use defense is just that: a defense that may prevail *notwithstanding* the presence of likely confusion. First, the burden of demonstrating confusion rests with the plaintiff; second, the terms of the defense itself are silent as to the relevance of likely confusion. "Starting from these textual fixed points, it takes a long stretch to claim that a defense of fair use entails any

---

> will inevitably destroy the advertising value of the mark." H.R. REP. NO. 104-374 (1995), *reprinted in* 1996 U.S.C.C.A.N. 1029, 1032.
>
> This passage is important in two respects. First, it evinces an intent to provide a broad remedy for the lesser trademark violation of dilution and recognizes that the essence of the dilution claim is a property right in the "potency" of a mark. While this does not reach the "property right in gross" proportions of Schechter's early dilution analysis, it does demonstrate an understanding that the right to be protected is in a mark's distinctiveness. Second, the passage's latter half—"confusion leads to immediate injury, while dilution is an infection, *which if allowed to spread*, will inevitably destroy the advertising value of the mark"—evinces an intent to allow a remedy *before* dilution has actually caused economic harm to the senior mark.

Id. at 475-76.

179. Nabisco, Inc. v. PF Brands, Inc., 191 F.3d 208, 224 (2d Cir. 1999).

180. *Id.* ("[S]uch a reading depends on excessive literalism to defeat the intent of the statute. Notwithstanding the use of the present tense in 'causes dilution,' it seems plausibly within Congress's meaning to understand the statute as intending to provide for an injunction to prevent the harm before it occurs."). The problem for the court was that the statute did not provide for damages unless the dilution was willful, creating the prospect that a trademark holder could not stop dilution until the harm had been consummated. *Id.* At that point, damages for the harm would likely be unavailable. *Id.*

181. 543 U.S. 111 (2004). Once again, notwithstanding the apparent simplicity of the matter as an exercise in textual interpretation, the case resolved a split in circuit authority. *Id.* at 116-17.

burden to negate confusion."[182] The text of the Act thus preserved a long-standing trademark defense from being wholly subsumed in the likelihood-of-confusion analysis.[183]

2. *Trademark "Contextualism"*

Another purported hallmark of today's formalism is the effort to rationalize bodies of law into coherent wholes, which requires statutory interpretation to fit within the larger context of the relevant field.[184] This goal dovetails with the modern textualist emphasis on context in statutory interpretation.[185] In trademark law, the contextualizing urge requires that

---

182. *Id.* at 118. The Court elaborated:

> It is just not plausible that Congress would have used the descriptive phrase "likely to cause confusion, or to cause mistake, or to deceive" in § 1114 to describe the requirement that a markholder show likelihood of consumer confusion, but would have relied on the phrase "used fairly" in § 1115(b)(4) in a fit of terse drafting meant to place a defendant under a burden to negate confusion. Where Congress includes particular language in one section of a statute but omits it in another section of the same Act, it is generally presumed that Congress acts intentionally and purposely in the disparate inclusion or exclusion. *Id.* (citations omitted) (alteration omitted) (quotation omitted). By the same token, the opinion turned away the senior user's attempt to argue that the defense's presence in the statute was akin to a drafting error and should have been removed when Congress amended the Lanham Act in 1989.

Id. at 120-21.

Justice Souter's opinion added two dollops of legislative history in footnotes (not joined by the ever-vigilant Justice Scalia), but they did not guide the analysis. *See id.* at 118 n.4, 122 n.5.

183. This is not to say that the ruling was an unqualified triumph for opponents of expansive trademark. The Court held open the door for the argument that the presence of likely confusion may affect a court's ruling on the viability of the fair use defense. *Id.* at 123 ("It suffices to realize that our holding that fair use can occur along with some degree of confusion does not foreclose the relevance of the extent of any likely consumer confusion in assessing whether a defendant's use is objectively fair."). On remand the Ninth Circuit took up the Court's invitation to dilute the effectiveness of the defense and held that likelihood of confusion is relevant to successfully establishing the defense. *See KP Permanent Make-Up*, 408 F.3d at 609.

184. *See supra* note 148 and accompanying text.

185. *See supra* note 147 and accompanying text; *see also* Eskridge, *supra* note 144, at 655, stating:

> Justice Scalia admits "coherence" arguments, that is, arguments that an ambiguous term is rendered clear if one possible definition is more coherent with the relevant legal authorities than other possible definitions. But, unlike defenders of legislative history, Justice Scalia admits only arguments based upon textual, or horizontal, coherence (this meaning is consistent with other parts of the statute or other terms in similar statutes), and not based upon historical, or vertical, coherence (this mean-

Lanham Act interpretations cohere with the rest of the statute as well as other realms of intellectual property law. The Supreme Court's treatment of functionality, origin, and product design as a form of trade dress reflects the desire for contextual coherence.[186]

a) Functionality

TrafFix Devices, Inc. v. Marketing Displays, Inc. offers a straightforward example of the Court's "trademark contextualism."[187] The Sixth Circuit rejected a district court ruling that the dual-spring design of a roadside sign (which prevented wind from blowing it down) was functional and therefore not protectable under section 43(a) of the Lanham Act.[188] The court of appeals believed that the district court should have considered the availability of alternatives to a dual-spring mechanism (e.g., a tri- or quad-spring design). A unanimous Court reversed. Once a product design is deemed functional, there is no need to engage in ad hoc consideration of the availability of alternatives.[189] Of note for present purposes, Justice Kennedy's opinion relied in part on the need to restrict trademark and patent law to their respective realms. Once a patent expires, the integrity of the patent regime requires free copying of the invention. Expansive trade dress protection would interfere with the time-limited monopoly bargain at

---

ing is consistent with the historical expectations of the authors of the statute).

186. *Cf. supra* note 148. As such, my use of the term "contextualism" means something quite different than that of Dinwoodie and Janis in their discussion of the trademark use doctrine. *See* Dinwoodie & Janis, *Lessons*, *supra* note 14, at 1708-09.

187. 532 U.S. 23 (2000).

188. *Id.* at 27. The Court stated:

> It was not sufficient, according to the Court of Appeals, that allowing exclusive use of a particular feature such as the dual-spring design in the guise of trade dress would "hinde[r] competition somewhat." Rather, "[e]xclusive use of a feature must 'put competitors at a *significant* non-reputation-related disadvantage' before trade dress protection is denied on functionality grounds."

Id. at 27-28 (alteration in original) (citations omitted).

The functional aspect of the springs was that they allowed the signs to yield to the wind without toppling over. An expired utility patent covered the design in question. *Id.* at 25.

189. *Id.* at 33-34 ("Here, the functionality of the spring design means that competitors need not explore whether other spring juxtapositions might be used. The dual-spring design is not an arbitrary flourish in the configuration of MDI's product; it is the reason the device works. Other designs need not be attempted."); *cf.* discussion *supra* note 148. Some courts continue to consider the existence of alternatives in deciding whether a design is functional in the first instance. *See, e.g.*, Valu Eng'g, Inc. v. Rexnord Corp., 278 F.3d 1268, 1276 (Fed. Cir. 2002).

the heart of patent law.[190]

b) The Meaning of "Origin"

Dastar Corp. v. Twentieth Century Fox Film Corp. demonstrates both the textualist preference for dictionaries over legislative history and the modern textualist emphasis on context.[191] *Dastar* involved a video producer who copied and edited an out-of-copyright television series and sold the resulting work under a new title.[192] Plaintiffs claimed that selling the series without attribution constituted "reverse passing off," causing likely consumer confusion as to the product's origin.[193]

Justice Scalia's analysis emphasized at the outset that the Lanham Act is not an open-ended common law cause of action to which judges may supply content. " '[B]ecause of its inherently limited wording, § 43(a) can never be a federal "codification" of the overall law of "unfair competition," ' but can apply only to certain unfair trade practices prohibited by its text."[194] The question, therefore, was the meaning of the word "origin" in section 43(a).[195] If "origin" simply means physical source, then the defendant was indeed the source of the product in question, but if "origin" means "author" or the like, then the plaintiffs might have a claim.[196]

---

190. *TrafFix*, 532 U.S. at 29 ("Trade dress protection must subsist with the recognition that in many instances there is no prohibition against copying goods and products. In general, unless an intellectual property right such as a patent or copyright protects an item, it will be subject to copying."). The Court stopped short of holding that the subject of an expired utility patent could *never* be protectable trade dress, *id.* at 35, but it was definitive that the existence of a prior patent was strongly probative of functionality, *id.* at 29-30, 32. Moreover, the Court's desire to keep the various forms of intellectual property protection to their proper domains is hardly a new concern. *See, e.g.*, Baker v. Selden, 101 U.S. 99, 102-03 (1879).

191. 539 U.S. 23 (2003). *Dastar* was an 8-0 opinion with Justice Breyer not participating.

192. *Id.* at 26-27.

193. *Id.* at 27. "Passing off" involves representing your product as that of another (e.g., putting a TOYOTA label on your homemade car and selling it as a Toyota). "Reverse passing off" arises when one takes the product of another and attempts to sell it under one's own mark (e.g., buying a Toyota and reselling it under the infringer's label).

194. *Id.* at 29 (quoting 4 MCCARTHY, *supra* note 26, § 27:7). Justice Scalia paused to note that courts had in the past arguably outrun their statutory authority. *Id.* at 29-30. In light of the Lanham Act's 1988 amendments, however, section 43(a)'s language is now "amply inclusive . . . of reverse passing off—if indeed it does not implicitly adopt the unanimous court-of-appeals jurisprudence on that subject."). *Id.* at 30.

195. *Id.* at 31.

196. The court explained:

> If "origin" refers only to the manufacturer or producer of the physical "goods" that are made available to the public (in this case the video-

Beginning with the dictionary definition of "origin," the Court concluded that "the most natural understanding of the 'origin' of 'goods'—the source of wares—is the producer of the tangible product sold in the marketplace."[197] While that concept may "stretch" to encompass a mark holder who ordered or otherwise "stood behind" the product, it cannot encompass the source of the intellectual content (e.g., the author or inventor) of the product.[198]

The Court's analysis did not end with the dictionary. It looked to broader context to interpret the term "origin."[199] The Court made a potentially challengeable empirical observation with respect to consumer expectations. In the Court's analysis, a purchaser of a Coke cares that her Coke will taste like the others she has tried, and is comparatively unconcerned with who invented the soda's formula in the first place.[200] This move, however, is fully consistent with the traditional common law understanding: trademark law has not traditionally focused on matters of authorial source.

More importantly, *Dastar* concedes that literary and similar works may be an exception and that some consumers might care about proper

---

> tapes), Dastar was the origin. If, however, "origin" includes the creator of the underlying work that Dastar copied, then someone else (perhaps Fox) was the origin of Dastar's product. At bottom, we must decide what § 43(a)(1)(A) of the Lanham Act means by the "origin" of "goods."

Id.

197. *Id.*

198. *Id.* at 31-32. Professor Dinwoodie argues that the text underdetermines the result in *Dastar*:

> But given that the language of Section 43(a) is clearly susceptible to more than one interpretation, one might suspect that there is something else going on. The *Dastar* Court appears willing to reject the endorsement of judicial development of this cause of action in the legislative history to the 1988 Berne Convention Implementation Act. Yet, the Court was also ready to accept in *Two Pesos* the endorsement of judicial expansion of the scope of trade dress actions in the legislative history to the 1988 Trademark Law Revision Act. Thus, mere statutory interpretation tools do not provide a complete explanation . . . .

Dinwoodie, *supra* note 173, at 203. A textualist might reply that the apparent contradiction is easily reconciled—legislative history does not matter. Congressional acquiescence to judicial practice as manifest in actual statutory text does. In any case, insofar as the differing results suggest a change in attitude toward "judicial development" of the trademark cause of action between 1992 and 2003, that supports the proposition that the Court's trademark jurisprudence jibes with the formalist narrative described above.

199. As modern textualist partisans maintain is appropriate. *See supra* note 147.

200. *Dastar*, 539 U.S. at 32.

attribution of authorship.[201] But to support such expectations would be to bring trademark into conflict with copyright law by allowing authors and their assignees to police activity that the Copyright Act permits.[202] In this *Dastar* is very much a contextualist ruling, drawing upon the boundaries set by other bodies of intellectual property law, as well as the common law foundations of trademark law, to interpret the Lanham Act's text.[203]

c) What About *Wal-Mart*?

Wal-Mart Stores, Inc. v. Samara Bros., Inc. presents a harder case for the claim that the recent history of Supreme Court trademark cases is strongly formalist, to say nothing of the larger thesis that today's courts are constrained in devising defenses to trademark liability.[204] In another unanimous opinion written by Justice Scalia, the Court held that unregistered product design is never inherently distinctive. Trade dress protection for product design always requires secondary meaning.[205]

Wal-Mart may seem a functionalist departure from the formalist cases surveyed thus far, particularly in light of the opinion's frequent references to policy considerations.[206] No text in the Act distinguishes between prod-

---

201. Professor Dinwoodie argues that this move is "amateur psychology" that presages what is actually a pragmatic opinion. Dinwoodie, *supra* note 173, at 204. It should be noted, however, that the opinion makes these observations *after* arriving at an interpretation of the semantic meaning of the term "origin" in the statute. The psychology, such as it is, is considered in determining whether to vary from the Court's statutory interpretation; it is not the basis for it.

"Amateur psychology" is fairly common in trademark law. See, e.g., Virgin Enters. Ltd. v. Nawab, 335 F.3d 141, 148 (2d Cir. 2003). ("[T]he more distinctive the mark, the greater the likelihood that the public, seeing it used a second time, will assume that the second use comes from the same source as the first."); Rogers v. Grimaldi, 875 F.2d 994, 1000 (2d Cir. 1989) ("[M]ost consumers are well aware that they cannot judge a book solely by its title any more than by its cover. We therefore need not interpret the Act to require that authors select titles that unambiguously describe what the work is about . . . .").

202. *Dastar*, 539 U.S. at 33. To hold otherwise would be to "create a species of mutant copyright law that limits the public's federal right to copy and to use expired copyrights." *Id.* at 34 (citation omitted) (quotation omitted). Moreover, identification of authors would present difficult challenges for Lanham Act purposes. *See id.* at 35-36.

203. *Id.* at 37 ("[R]eading the phrase 'origin of goods' in the Lanham Act in accordance with the Act's common-law-foundations (which were *not* designed to protect originality or creativity), and in light of the copyright and patent laws (which *were*), we conclude that the phrase refers to the producer of the tangible goods[.]").

204. Wal-Mart Stores, Inc. v. Samara Bros., Inc., 529 U.S. 205, 211 (2000).

205. *Id.* at 212.

206. *See, e.g.*, *id.* at 214 (commenting on virtue of "summary dispositions of an anticompetitive strike suit"); *id.* ("Competition is deterred, however, not merely by successful suit but by the plausible threat of successful suit . . . ."); *id.* at 213-14 (noting *in terro-*

uct design and packaging. Is this then a case of the Court's inventing doctrine to blunt the Lanham Act's extremes? No.

Wal-Mart is a contextualist holding in the same vein as others canvassed in this Section. It is textualist insofar as it elucidates statutory text based on the context provided by the rest of the statute. The resulting rule is an elaboration of the statute itself, not a product of common law statutory updating. The problem for the Court was that Congress updated the Lanham Act to broaden its cause of action, but omitted many important details of how the section 43(a) cause of action was to function. What may function as an unregistered trade dress was one such open issue.

Notwithstanding this silence, *Wal-Mart*'s analysis still begins with the Lanham Act's text, specifically the description of what may be a *registered* trademark.[207] The opinion notes the contestable interpretive move made over time by lower courts to allow protectable trade dress to include product design, but cites *Qualitex*'s interpretation of the terms "symbol" and "device" in trademark's definition to support permitting the practice.[208]

This says nothing, however, about what sort of *unregistered* trade dress is protectable.[209] To address this issue, *Wal-Mart* looks first not to the common law or the Court's perception of the best policy, but rather to the statutory standards that exist for registered marks.[210] Those standards are found in section two of the Lanham Act,[211] which reflects the longstanding distinction between inherently distinctive marks, which are automatically eligible for protection, and non-distinctive marks, which receive protection only if consumers have come to associate them with a single source.[212] This contextualist move leaves open an important ques-

---

rem effect of uncertainty to competitors); *id.* at 215 (noting that balance of utilities favors the Court's preferred rule).

207. *Id.* at 209 (discussing section 2 of the Lanham Act).

208. *Id.* Similarly, Congress appeared to acquiesce in the judicial practice of protecting product design under section 43(a) by amending the statute to refer specifically to such actions. *Id.* (noting § 43(a)(3), which places burden on the party claiming protection in unregistered trade dress to establish its lack of functionality).

209. *Id.* at 209-10 ("The text of § 43(a) provides little guidance as to the circumstances under which unregistered trade dress may be protected.").

210. Looking to standards for registered marks to fill the content of the cause of action for unregistered marks is commonplace for actions arising under section 43(a), *see* 5 MCCARTHY, *supra* note 26, § 27:18, and was the Court's method in *Two Pesos* as well. *See supra* note 37 and accompanying text.

211. 15 U.S.C. § 1052 (2006) (setting forth registration requirements).

212. *Wal-Mart*, 529 U.S. at 210. The Court explained:

> The judicial differentiation between marks that are inherently distinctive and those that have developed secondary meaning has solid foun-

tion: does every category of mark contain inherently distinctive marks? "Nothing in § 2 . . . demands the conclusion."[213]

Wal-Mart thus identifies the applicable rule—no protection for non-inherently distinctive marks absent secondary meaning—and an open question: Is product design ever inherently distinctive? It is only once the text runs out that the Court crafts its prophylactic rule, one giving breathing space to the other relevant textual command of the Lanham Act: functional subject matter does not get trademark protection.[214] Just as *Dastar* sought to prevent an innovative trademark claim from intruding on copyright's realm, so *Wal-Mart* prevents a newer form of mark protection from encroaching upon a long-standing exclusion. In short, *Wal-Mart* is not fairly described as a "common law" trademark case. Rather, it is statutory construction of the elements of the Lanham Act's cause of action. *Wal-Mart* implements Act's commands of what may be a mark in the first place.

Perhaps none of this is persuasive.[215] One may argue that *Wal-Mart* is little more than a judgment rooted largely in another contestable empirical claim about how consumers behave. The same might be said, however, about a ruling to the contrary. Concluding that consumers (or some of them) see product design as inherently distinctive is also a factual judgment. Admittedly, this uncertainty may be reason enough to refer the decision to the factfinder on a case-by-case basis (notwithstanding considerations of judicial economy cited by the Court).[216]

> dation in the statute itself. Section 2 requires that registration be granted to any trademark "by which the goods of the applicant may be distinguished from the goods of others"—subject to various limited exceptions. 15 U.S.C. § 1052. It also provides, again with limited exceptions, that "nothing in this chapter shall prevent the registration of a mark used by the applicant which has become distinctive of the applicant's goods in commerce"—that is, which is not inherently distinctive but has become so only through secondary meaning. § 2(f), 15 U.S.C. § 1052(f).

Id. at 211.

213. *Id.*

214. 15 U.S.C. §§ 1052(e)(5), 1115(b)(8), 1125(a)(3) (2006).

215. For his part, Professor Dinwoodie views *Wal-Mart* as being equally grounded in concern for dangers to competition as in the text of the Lanham Act or the principles of trademark law. Dinwoodie, *supra* note 173, at 198 (labeling *Wal-Mart*'s holding as a "prudentially derived conclusion"). *Cf.* Merrill, *supra* note 152, at 372 (critiquing textualism as seemingly "transform[ing] statutory interpretation into a kind of exercise in judicial ingenuity").

216. Professor Dinwoodie argues that the Court could have implemented its conclusions regarding the relative source identifying capabilities of product design and product

Note, however, that individualized scrutiny of the purported trade dress's viability remains possible under *Wal-Mart*. The analysis simply must proceed through the prism of secondary meaning, a form of analysis rooted in the common law of unfair competition and the text of the Lanham Act.[217] The *Wal-Mart* rule, therefore, channels the necessary case-by-case determinations into a framework familiar to trademark law rather than relying on ad hoc standards (which may or may not respect the functionality bar) to separate inherently distinctive from non-inherently distinctive designs.[218] *Wal-Mart* thus fits the new formalist narrative that favors enhancing predictability by trying to bring internal coherence to bodies of law and minimizing the unpredictability of case-by-case adjudication.[219]

What is important is that the Court viewed itself as doing something more modest than crafting a new common law rule to ride atop existing trademark doctrine. Whatever its faults, the dress/design distinction is an elaboration of the inherent distinctiveness requirement of section 2 of the Lanham Act. In other words, the Court found its freedom in the Lanham Act's text and structure—not, as might appear at first glance, in the statute's silences.[220] Contextual application of open-ended text is statutory

---

packaging by crafting a rule that would still allow courts to evaluate actual consumer perceptions on a case-by-case basis. "Instead, the Court foreclosed individualized scrutiny of its (unsupported) social generalization, by embedding that generalization as a rule of law." Dinwoodie, *supra* note 173, at 197.

217. 15 U.S.C. § 1052(f) ("Except as expressly excluded . . . nothing in this chapter shall prevent the registration of a mark used by the applicant which has become distinctive of the applicant's goods in commerce.").

218. This is not to say, of course, that the secondary meaning inquiry is always predictable, simply that it is familiar to trademark law and finds a textual basis in statute. No similarly deep or long-lived body of law had evolved regarding the inherent distinctiveness of product design. *Compare* 1 SHOEMAKER, *supra* note 34, at 208 (describing secondary meaning doctrine as understood in 1931), *with id.* § 78 at 236 ("A trade-mark must be something distinct from the article marked; neither the thing itself nor any part or quality of it can be a trade-mark for that thing.").

219. *See supra* note 148 and accompanying text.

220. In a similar vein, *Burlington Indus., Inc. v. Ellerth*, 524 U.S. 742 (1998) created an affirmative defense to employer liability under Title VII based on the statute's definition of "employer" as including the term "agents." The Court stated, "Congress has directed federal courts to interpret Title VII based on agency principles. Given such an explicit instruction, we conclude a uniform and predictable standard must be established as a matter of federal law." *Id.* at 754. Quoting *Meritor Sav. Bank, FSB v. Vinson*, 477 U.S. 57, 72 (1986), the Court concluded that "Congress' decision to define 'employer' to include any 'agent' of an employer, 42 U.S.C. § 2000e(b), surely evinces an intent to place some limits on the acts of employees for which employers under Title VII are to be held responsible." *Id.* at 763.

interpretation and construction, not an exercise in common law jurisprudence.

C. Summary

While the contention that we are in a formalist era may be contestable, the Supreme Court's recent approach to trademark cases is consistent with the claim. This may be bad news for trademark defenses. While the opinions canvassed above generally restrict trademark's scope for textualist reasons, the text of the Lanham Act generally supports today's expansive trademark doctrines. The real bite of the Court's "trademark formalism" may therefore be felt as an obstacle to future efforts at reform.

In *Wal-Mart*, however, the Court revealed one promising avenue for future development. The Court framed its holding as statutory construction of open-ended text. Those looking to create similar pro-defendant innovations might benefit from a search for similar examples of open text in the Lanham Act.

IV. WHAT'S LEFT FOR TRADEMARK DEFENSES?

This Part surveys available sources of future innovation in trademark defenses. It concludes with a discussion of the "nominative fair use" doctrine as an example of the uncertainty and problems facing courts that would craft new defenses.

A. The Source of Trademark Defenses

In considering trademark defenses, judges may look to the Lanham Act's text, the "federal common law" of trademark, and external legal requirements. This Section examines why none of these sources are a fully adequate basis for future judicial creativity.

1. The "Literal" Lanham Act

Section 33(b) of the Lanham Act codifies several traditional common law defenses to trademark infringement and applies them to suits for the infringement of registered marks, regardless of incontestable status.[221] The

---

221. After five years of use, a holder of a registered mark may obtain incontestable status. 15 U.S.C. § 1065 (2006). Incontestable status is "conclusive evidence of the validity of the registered mark and of the registration of the mark, of the registrant's ownership of the mark, and of the registrant's exclusive right to use the registered mark in commerce." 15 U.S.C. § 1115(b) (2006).

Section 33(b) of the Lanham Act sets forth a list of defenses and defects to which a mark is subject notwithstanding incontestable status. 15 U.S.C. § 1115(b). Professor McCarthy argues that the listed defenses and defects were intended by Congress to

enumeration of specific defenses has always raised the question of whether the list is exclusive with respect to incontestable marks.[222] The Court has not considered the issue lately,[223] but the most recent relevant precedent suggests a closed list.

Park 'N Fly, Inc. v. Dollar Park & Fly, Inc.[224] considered whether an alleged infringer may defend on the basis that the plaintiff's incontestable mark was nonetheless descriptive, and thus ineligible for protection.[225] The answer was simple because section 33(b) contains no exceptions for descriptive marks.[226] To hold otherwise would "emasculate[]" the relevant statutory provision.[227]

Park 'N Fly implies that courts have scant room to maneuver with respect to defenses for infringing incontestable marks.[228] Partisans of tex-

---

"merely reduc[e] the status of a conclusive presumption down to that of prima facie, with the challenger allowed to raise common law defenses," notwithstanding the judicial practice of reading the provision as setting forth defenses on the merits. 6 MCCARTHY, *supra* note 26, § 32:157.

222. *Compare* Diggins, *supra* note 59, at 195 ("The fact that Section 33(b) limits the defenses against an incontestable mark to seven specific issues is possibly not conclusive. It is difficult to imagine an equity court granting injunctive relief to a registrant who comes into court with unclean hands . . . ."), *with* Symposium, *Incontestable Trademark Rights and Equitable Defenses in Infringement Litigation*, 66 MINN. L. REV. 1067 (1982) (arguing that equitable defenses were unavailable under then-current Lanham Act text, which omitted the defenses). Congress resolved the question in 1988 by adding equitable defenses to section 33(b). Trademark Law Revision Act of 1988, Pub. L. No. 100-667, sec. 128(b), § 33(b), 102 Stat. 3935, 3944 (current version at 15 U.S.C. § 1115(b) (2006)).

223. In *KP Permanent* the Court reserved comment on the role of alternative defenses like nominative fair use. KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc., 543 U.S. 111, 115 n.3 (2004). Nominative fair use is discussed in greater detail below. *See infra* Section IV.C.

224. 469 U.S. 189 (1985).

225. Because a descriptive mark should not be registered unless the registrant can show that secondary meaning had been established, *Park 'N Fly* covered the situation in which an incontestable trademark was arguably improperly registered in the first instance. *See* 15 U.S.C. § 1052(f) (2006).

226. *Park 'N Fly*, 469 U.S. at 196-97.

227. *Id.* at 197. Unlike the rather sparse use of legislative history in the Court's opinions discussed above, and consistent with the scholarship that reports declining use of such history, *see supra* note 152, Justice O'Connor's majority opinion included a section on the Lanham Act's legislative history and policies and argued that neither contradicted the clear dictates of the text. *Park 'N Fly*, 469 U.S. at 197-202.

228. *Cf.* Shakespeare Co. v. Silstar Corp. of Am., 9 F.3d 1091 (4th Cir. 1993) (holding incontestable mark may not be cancelled due to functionality of mark). *But compare* Wilhelm Pudenz, GmbH v. Littlefuse, Inc., 177 F.3d 1204, 1209 (11th Cir. 1999) (disagreeing with *Shakespeare*).

tualism might note that *Park 'N Fly* accomplished precisely what rigorous adherence to statutory provisions is supposed to do: force Congress to speak on ambiguous matters. Since *Park 'N Fly*, Congress has twice added new defenses to section 33(b), providing that an equitable defense and the defense of functionality may be raised in response to an infringement claim of an incontestable mark.[229] These actions, in turn, only reinforce the *inclusio unius* ramifications of opinions like *Park 'N Fly*. It cannot be argued that Congress has not considered the appropriate content of the provision. Barring further congressional action, section 33(b) seems a closed set, which looms as a problem for any defensive innovations that cannot fit within its provisions.[230]

---

229. Trademark Law Treaty Implementation Act, Pub. L. No. 105-330, sec. 201(a)(9), § 33(b), 112 Stat. 3064, 3070 (1998) (codified as amended at 15 U.S.C. § 1115(b) (2006)) (amending section to include functionality defense); Trademark Law Revision Act of 1988, Pub. L. No. 100-667, sec. 128(b)(5), (6), §§ 33(b), 34(a), 102 Stat. 3935, 3944-45 (codified as amended at 15 U.S.C. §§ 1115(b), 1116(a) (2006)) (amending section to include equitable defenses); *see also Wilhelm Pudenz*, 117 F.3d at 1211 ("[T]he legislative history of the Trademark Law Treaty Implementation Act indicates that the functionality provisions were meant to codify existing law and correct the flawed result reached by the Fourth Circuit in *Shakespeare*.").

230. Nor does the availability of "equitable principles" under section 33(b)(9) give courts room for innovation. In context, it is clear that the codified equitable defenses are the traditional ones, as the defense is implicated when "equitable principles, *including laches, estoppel, and acquiescence*, are applicable." 15 U.S.C. § 1115(b)(9) (emphasis added). This is the necessary implication of *Park 'N Fly*, which held that the provision authorizing injunctions "according to the principles of equity" did not open the door to asserting defenses based on a mark's descriptiveness. *Park 'N Fly*, 469 U.S. at 203. As the Court explained:

> Whatever the precise boundaries of the courts' equitable power, we do not believe that it encompasses a substantive challenge to the validity of an incontestable mark on the grounds that it lacks secondary meaning. To conclude otherwise would expand the meaning of "equity" to the point of vitiating the more specific provisions of the Lanham Act.

Id.

In addition, the Court reserved the question of whether "traditional equitable defenses such as estoppel or laches" apply (as the case arose prior to the inclusion of defenses in section 33(b)). *Id.* at 203 n.7. If such equitable defenses were capable of extending beyond traditional equitable doctrines, the Court would have had to address the matter or explain why not (e.g., due to a party's waiver). The legislative history of section 33(b)(9) is in accord. S. REP. NO. 100-515 at 39 (1988), *as reprinted in* 1988 U.S.C.C.A.N. 5577; *see also* 6 MCCARTHY, *supra* note 26, § 32:151 ("The ability to raise 'equitable principles' does not open the door to any and all defenses. It is not a catch-all category." (citing Levi Strauss & Co. v. GTFM, Inc., 196 F. Supp. 2d 971 (N.D. Cal. 2002))).

2. *"Federal Common Law" Trademark Defenses*

While the express defenses of the Lanham Act are limited, they are not the end of the story. Section 33(a) signals the existence of other defenses,[231] but says nothing about what they are, nor does it contain language authorizing courts to innovate.[232] The statute is likewise silent with

---

231. 15 U.S.C. § 1115(a) (2006) (registration "shall not preclude another person from proving any legal or equitable defense or defect, including those set forth in subsection (b) of this section, which might have been asserted if such mark had not been registered").

232. It is therefore best read as incorporating trademark requirements from elsewhere in the statute and allocating, where necessary, the burden of going forward. So, for example, a registered (but not incontestable) mark may be presumed valid, but a defendant may still claim that it is defective because it is descriptive. *See* 15 U.S.C. § 1115(a) (registration is prima facie evidence of validity), *cf. Park 'N Fly*, 469 U.S. at 195-97. *Park 'N Fly* is in accord:

> The Lanham Act expressly provides that before a mark becomes incontestable an opposing party may prove any legal or equitable defense which might have been asserted if the mark had not been registered. Thus, § 33(a) would have allowed respondent to challenge petitioner's mark as merely descriptive if the mark had not become incontestable.

Id. at 196 (citation omitted).

Judge Leval has argued to the contrary, contending that the Lanham Act is a delegating statute that stands for "complete common law development," and that section 33(a)'s language is in fact an explicit delegation to the courts. Pierre N. Leval, *Trademark: Champion of Free Speech*, 27 COLUM. J.L. & ARTS 187, 198 (2004); *see also* Dinwoodie, *Developing Defenses*, *supra* note 13, at 138. While the open text of the statute authorizes some judicial creativity, *see infra* Section IV.B, the clause in question is not plausibly read as doing so. In full, the surrounding sentence reads:

> Any registration issued under the Act of March 3, 1881, or the Act of February 20, 1905, or of a mark registered on the principal register provided by this chapter and owned by a party to an action shall be admissible in evidence and shall be prima facie evidence of the validity of the registered mark and of the registration of the mark, of the registrant's ownership of the mark, and of the registrant's exclusive right to use the registered mark in commerce on or in connection with the goods or services specified in the registration subject to any conditions or limitations stated therein, but shall not preclude another person from proving any legal or equitable defense or defect, including those set forth in subsection (b) of this section, which might have been asserted if such mark had not been registered.

15 U.S.C. § 1115(a). In other words, the effect of the clause is to limit the power of registration with respect to existing defenses or defects. Holders of registered marks are entitled to certain presumptions, but their rights are subject to other provisions of trademark law. A caveat to a provision providing benefits to trademark registrants seems an odd place to bury broad authorization to judges to craft new common law defenses.

Even if the language is read as authorizing interstitial lawmaking, it remains another leap to interpret it as inviting judges to go beyond trademark's existing common law

respect to defenses to infringement of unregistered marks under section 43(a). Nonetheless, in what has been described as an exercise of the federal common law, courts apply the defenses of section 33(b) when adjudicating cases under section 43(a).[233] Why then may courts not go further and craft new defenses as part of the interstitial law of section 43(a)?[234]

First, past practice applying section 33(b) to section 43(a) cases may not be as "common law" as it appears.[235] In much the same manner that the requirements for trademark registration inform the scope of protection for unregistered marks, so the codified defenses for infringement of incontestable marks inform the scope of the rights held by a holder of an unregistered mark. In other words, courts are arguably engaging in contextualist statutory construction like that in *Wal-Mart*.[236] If so, it bears noting that *Wal-Mart* adopted a rule effectuating a doctrine found in the Lanham Act's text; it did not go beyond the statutory principle.

Another way to approach the problem is to note that Congress wrote against existing background principles of law when it ratified the judicial expansion of section 43(a). The resulting handiwork should therefore be understood and interpreted against that legal context.[237] In other words,

---

backdrop. *See infra* notes 243-248 and accompanying text; *cf. Dastar*, 539 U.S. at 29 ("[B]ecause of its inherently limited wording, § 43(a) can never be a federal 'codification' of the overall law of 'unfair competition,' but can apply only to certain unfair trade practices prohibited by its text.") (citation omitted) (internal quotations omitted).

233. 5 MCCARTHY, *supra* note 26, § 27:19 ("[T]he statutory 'defenses' in a § 43(a) case are merely guidelines to ascertain the federal common law substantive 'defenses' to a § 43(a) claim.").

234. Note, however, that even if discretion exists to innovate under section 43(a), the holding in *Park 'N Fly* suggests that such innovations cannot easily migrate to the cause of action for infringement of registered marks under section 32.

235. The term "common law" is used in the above sentence in the federal common law sense of interstitial lawmaking to fill gaps left by Congress. *See generally* ERWIN CHEMERINSKY, FEDERAL JURISDICTION 355-56, 376 (4th ed. 2003) ("Federal common law has developed out of necessity. In some instances there are simply gaps in the law; the application of statutory and constitutional provisions often requires the development of legal rules.") (footnote omitted).

236. *See supra* Section III.B.2.

237. John F. Manning, *The Absurdity Doctrine*, 116 HARV. L. REV. 2387, 2467 (2003) [hereinafter, Manning, *Absurdity*] ("If the meaning of a text depends on the shared background conventions of the relevant linguistic community, then any reasonable user of language must know 'the assumptions shared by the speakers and the intended audience.' " (quoting Frank H. Easterbrook, *What Does Legislative History Tell Us?*, 66 CHI. KENT L. REV. 441, 443 (1991))). In this manner, a number of doctrines are routinely applied to text that does not explicitly invoke them (even if those doctrines were the product of judicial intervention before becoming part of the legal background against which future legislatures acted). *Id.* at 2466-70 (collecting examples).

courts interpret the Lanham Act with long-standing common law and statutory defenses in mind, and read open statutory text accordingly.[238] This interpretation constrains rather than liberates the courts. It is one thing to assume that preexisting practices survive passage of a statute that does not explicitly negate them.[239] It is quite another to assume authorization to create new defenses that lack any statutory tether. Such an assertion rests on an entirely different conception of the judicial power.[240] Absent a textual basis, there are no standards for courts to employ. While that may not have mattered in the past, it does today.[241]

---

238. Manning, *What Divides?*, *supra* note 146, at 82 & n.42 (observing that "settled" common law practices may be part of unstated background in which statute is understood to act and citing equitable tolling of statutes of limitations and interpretations of criminal statutes as examples); John F. Manning, *Textualism and Legislative Intent*, 91 VA. L. REV. 419, 435-36 (2005) ("Textualists assign common-law terms their full array of common-law connotations; they supplement otherwise unqualified texts with settled common-law practices, where such practices traditionally pertained to the subject matters covered within the statute . . . .").

This reflects the practice of the courts with respect to incorporation of common law doctrines of geographic scope. Acting in the early part of the Twentieth Century, the Supreme Court had restricted the geographic scope of trademark rights to the active markets and zones of natural expansion of markholders. *See generally* 5 MCCARTHY, *supra* note 26, § 26:1–:30 (discussing common law use rights). This doctrine persists with respect to the geographic scope of unregistered marks enforceable by section 43(a). *Id.* § 26:52 (describing persistence of doctrine as federal common law).

239. Following this logic, *Wilhelm Pudenz, GmbH v. Littlefuse, Inc.* held that incontestable marks may be canceled on functionality grounds even though the Lanham Act did not mention the doctrine at the time the litigation arose:

> Consequently, the mere fact that functionality is not enumerated in § 1115(b) is not sufficient to indicate congressional intent to eliminate the defense's applicability to incontestable registrations. Indeed, given the absence of any explicit reference to the functionality doctrine, which is a judicially created concept that predates the Lanham Act, we should be hesitant to read the Act as limiting the doctrine's reach. "The normal rule of statutory construction is that if Congress intends for legislation to change the interpretation of a judicially created concept, it makes that intent specific."

177 F.3d 1204, 1210 (11th Cir. 1999) (quoting *Midlantic Nat'l Bank v. New Jersey Dep't of Envtl Prot.*, 474 U.S. 494, 501 (1986)).

240. Manning, *Absurdity*, *supra* note 237, at 2466 ("Modern textualists unflinchingly rely on legal conventions that instruct courts, in recurrent circumstances, to supplement the bare text with *established* qualifications designed to advance certain substantive policies.") (emphasis added).

241. *See supra* Section III.B. And even if legislative history bears on the question, the committee report on the amendment of section 43(a) provides no call for defensive innovation. In announcing a codification of past broad interpretations of liability under section 43(a), the report notes that "[b]ecause Section 43(a) of the Act fills an important gap

Even if one concedes some "gap filling" authority on the part of the courts (based either on the absence of detail in section 43(a) or the conundrums raised by giving unregistered trademarks greater protection than their registered counterparts),[242] that does not necessarily invite further innovation. The practice of looking to section 33(b), whether as a matter of construction or interstitial lawmaking, is a text-bound, rather than open ended, inquiry. It therefore suggests a *restrictive* rather than an expansive view of what federal "common law" defenses may be.

Another problem with the defense-as-interstitial-lawmaking argument, leaving aside the lack of clear statutory authorization,[243] is that it misconceives the nature of the "gap" left by Congress. It may be true that the federal trademark cause of action demands application of a federal rule, rather than a borrowed state rule, to a question arising within the "policy bundle" of the cause of action.[244] That does not mean that the judge has

---

in federal unfair competition law, the committee expects the courts to continue to interpret the section." S. REP. NO. 100-515 (1988), *as reprinted in* 1988 U.S.C.C.A.N. 5577, 5603. If anything, the statement in context cheers on further expansions of liability, not defensive carveouts.

Even the activist period of Lanham Act interpretation is the exception, not the rule in American jurisprudence. *See* John F. Manning, *Lawmaking Made Easy*, 10 GREEN BAG 2D 191, 196 (2007) (discussing Judge Friendly's endorsement of the view that "judges should treat congressional grants of jurisdiction over particular substantive areas as invitations to develop federal common law rules of decision" and observing that view "is hardly a pervasive feature of American public law"). One hardly needs to be a textualist to appreciate the costs to predictability and certainty of leaving judges with broad leeway to fill perceived statutory gaps. They are clear to pragmatists. POSNER, *supra* note 168, at 49 ("Moderate legalists are matched by moderate pragmatists—pragmatists who believe that the institutional consequences of judicial decisions argue for a judicial approach heavily seasoned with respect for the language of contracts, statutes, and precedents."); *see also supra* note 159.

242. Potentially raising an absurdity objection, which remains an acceptable interpretive tool for textualist judges if not scholars. Manning, *Absurdity*, *supra* note 237, at 2391 (criticizing the doctrine from a textualist perspective, but observing that "even the staunchest modern textualists still embrace and apply, even if rarely, at least some version of the absurdity doctrine.").

243. CHEMERINSKY, *supra* note 235, § 6.3 at 379-80 ("The federal judiciary will formulate a body of common law rules only pursuant to clear congressional intent for such action.").

244. Caleb Nelson, *The Persistence of General Law*, 106 COLUM. L. REV. 503, 544 (2006). Nelson explains:

> Despite the variety of choice-of-law rules used in different American jurisdictions, those rules all tend to treat as a package certain issues that accompany the creation of a cause of action. Under virtually all American choice-of-law regimes, for instance, the same state's law that governs whether a cause of action exists (and what its elements are) will

unbounded discretion. As Caleb Nelson argues, judges fill such gaps with rules of "general law—rules whose content is not dictated entirely by any single decisionmaker (state or federal), but instead emerges from patterns followed in many different jurisdictions."[245] For example, the question of whether a federal cause of action survives the death of a party has frequently been decided by reference to evolving common law principles, rather than borrowing local law.[246]

As applied to the question of Lanham Act defenses, even though a federal court is unlikely to incorporate state law, its discretion with respect to section 43(a) is limited because application of the general law looks to existing defenses rather the judge's own ingenuity.[247] In other words, the existing common law backdrop of the trademark cause of action guides the judge. The resulting conservative presumption against innovation is only reinforced by Congress's specific incorporation of traditional, existing defenses in section 33(b) (statutory defenses to the infringement of incontestable marks). Worse, channeling trademark litigation to the federal courts limits the prospect that future defensive innovations might emerge from state common law.[248]

> also govern . . . the existence of substantive defenses . . . .

Id.

245. *Id.* at 503.

246. *Id.* at 545-46; *cf.* Hard Rock Cafe Licensing Corp. v. Concession Servs., Inc., 955 F.2d 1143, 1148 (7th Cir. 1992) (regarding questions of contributory trademark infringement, the courts "have treated trademark infringement as a species of tort and have turned to the common law to guide our inquiry into the appropriate boundaries of liability"). Not all Lanham Act questions have been seen as similarly part and parcel of the cause of action. *See, e.g.*, Sears, Roebuck & Co. v. Sears Realty Co., 932 F. Supp. 392, 401 (N.D.N.Y. 1996) (concluding that state law must govern the validity of settlement agreements under the Lanham Act "because there is no federal statute or common law rule on point that provides a rule of decision, and because the circumstances do not justify the creation of a federal common law rule"). But the general practice of gap filling under section 43(a) has been to look to either the standards for federally registered marks or traditional common law standards, not the invention of new law. 5 MCCARTHY, *supra* note 26, § 27:18.

247. Nelson, *supra* note 244, at 503 ("[W]hen courts articulate rules of 'federal common law' to fill vacuums created by written federal law, they assert less creative power than modern commentators typically suggest."); *cf. id.* at 548 ("[T]he basic rule is simple: Absent contrary guidance from Congress, statutes creating federal causes of action to enforce federal duties are typically understood not only to federalize questions about the proper measure of damages, but also to draw the substance of the federal rules from principles of general law.").

248. 4 MCCARTHY, *supra* note 26, § 23:1.50 (observing that courts generally apply a unified likelihood-of-confusion analysis to cases combining federal and state trademark claims and that unified approach allows citation of federal precedent); USTA Report,

The creation of new common law defenses faces a final, more conceptual problem. An affirmative defense is not any old statutory gap. A true defense defines an activity that is not subject to liability. To innovate in this area would be to take a congressional declaration that act X violates the law and declare that X does *not* violate the law. Beyond those defenses, like statute of limitations, that have always been seen as part and parcel of American jurisprudence (and thus the backdrop against which Congress legislated), defensive innovations seem to negate the statute itself.[249]

Nor is it a reply to argue that the broad *liability-creating* provisions of the Lanham Act compel a reciprocal flexibility in the availability of unenumerated defenses. If the chosen language of the Lanham Act controls in the absence of any manifest ambiguities, then we must distinguish situations in which Congress legislated narrowly, as in the case of enumerated defenses, from those in which it legislated broadly, as it did by choosing open-ended language for the law's causes of action.[250] Congress is free to legislate with varying levels of specificity, and the courts must live with the results.[251]

---

supra note 33, at 377 (observing "the strongly federal cast" of trademark law and policy and noting that "federal courts now decide, under federal law, all but a few trademark disputes. State trademark law and state courts are less influential than ever.").

249. This is especially so given that section 43(a), while containing open provisions, is hardly without standards. *Cf.* Tex. Indus., Inc. v. Radcliff Materials, Inc., 451 U.S. 630, 643-47 (1981) (concluding that while the Sherman Act authorizes creation of a federal common law, similar authorization does not appear on face of treble damages provision to create a right of contribution); Getty Petroleum Corp. v. Island Transp. Corp., 862 F.2d 10, 16 (2d Cir. 1988) (applying a similar conclusion to the Lanham Act).

250. In other words, some provisions of the Lanham Act may be characterized as delegating broad authority to the courts to implement Congress's wishes, but others are not open to that interpretation. The law's provisions on defenses fall into the latter category. *See supra* note 232 and accompanying text. Congress is capable of legislating open defenses in the intellectual property realm, as it did in codifying the fair use defense in copyright. 17 U.S.C. § 107 (2006).

251. *Cf.* Frank H. Easterbrook, *Text, History, and Structure in Statutory Interpretation*, 17 HARV. J.L. & PUB. POL'Y 61, 68 (1994) ("Sometimes Congress specifies values or ends, things for the executive and judicial branches to achieve, but often it specifies means, creating loopholes but greater certainty."); Manning, *What Divides?*, *supra* note 146, at 105 ("If interpreters pay attention to the way a reasonable person would understand language in context, then legislative drafters can choose to convey policy directives with greater or lesser degrees of specificity.").

As Judge Easterbrook argues, a legislature may pursue a goal by allowing courts to design rules or by designing a rule itself. While legislatively selected rules are "bound to be imprecise, to be over- and under-inclusive," this does not justify a judge's decision

> to add to or subtract from Rule Y on the argument that, by doing so, it

3. *External Constraints*

Trademark law is not an island, and external constraints—particularly those imposed by the First Amendment—need to be accommodated. While such accommodations may form the basis of defensive doctrines, their potential is limited.

We have encountered several analogous accommodations in our discussion of the Supreme Court's "contextualist" trademark rulings. Taking cues partly from the existence and requirements of copyright and patent law, the Court established a pair of defendant-friendly rules: (1) the Lanham Act does not address failures to attribute authorship of creative works; and (2) there is no protection for functional product design despite the availability of competitive alternatives. It should be noted that these external doctrines were not used as sources of the resulting rules, but merely guided the interpretation of provisions *internal* to the Lanham Act itself or the traditional law of trademark.[252]

The First Amendment is a potentially powerful source of trademark rules external to the Lanham Act. On one level, free speech interests often appear to exert a strong gravitational pull on trademark analysis.[253] On another, the fear of conflict with expressive rights has led various courts to create balancing tests to ensure that trademark liability maintains a safe distance from expressive considerations.[254]

That said, the commercial speech doctrine restricts the First Amendment's promise as a significant check to trademark's scope. The reduced scrutiny applied to regulation of trademark speech typically limits the

---

> can get more of Goal X. The judicial selection of means to pursue X displaces and directly overrides the legislative selection of ways to obtain X. It denies to legislatures the choice of creating or withholding gapfilling authority.

Frank H. Easterbrook, *Statutes' Domains*, 50 U. CHI. L. REV. 533, 546-47 (1983).

252. *See supra* Section III.B.2.

253. *E.g.*, Hormel Foods Corp. v. Jim Henson Prods., Inc., 73 F.3d 497, 503 (2d Cir. 1996) (considering parodic intent as part of finding no likely confusion between SPAM trademark and puppet named "Spa'am").

254. Rogers v. Grimaldi, 875 F.2d 994, 998 (2d Cir. 1989) ("Because overextension of Lanham Act restrictions in the area of titles might intrude on First Amendment values, we must construe the Act narrowly to avoid such a conflict."). In *Rogers*, the Second Circuit concluded that use of a trademark in an artistic work's title is infringing "only where the public interest in avoiding consumer confusion outweighs the public interest in free expression." *Id.* at 999. To that end, the court adopted a balancing test looking to whether the use of the mark is artistically relevant to the underlying work. If the first prong is met, the test asks whether the use explicitly misleads regarding source or content. *Id.*

room available for judges inclined to innovate in the defense area.[255] The more typical reaction of the courts is to treat the trademark realm as fenced off from speech concerns.[256]

B. What's Left? The "Implied" Lanham Act

Although judicial creativity with respect to defenses qua defenses may be limited, there is room to maneuver in other provisions of the Lanham Act. While the statute enumerates specific defenses, it is vaguer with respect to the standards that govern its substantive cause of action. In other words, the statute's open-ended liability-creating provisions may offer the best bet for the creation of de facto defensive doctrines.

To take a familiar example, before a trademark claim may succeed, a plaintiff must have a protectable mark. As discussed in Section III.B.2.c), existing law allows most anything to be a mark so long as it is distinctive. But it was the very openness of this requirement that enabled *Wal-Mart*'s distinction between product packaging (which may be inherently distinctive) and product design (which always requires secondary meaning). In other words, the defendant-friendly innovation stemmed from the Act's open provisions on mark validity and not its comparatively closed provisions on defenses.

The likelihood of confusion standard is the source of several similar judicial elaborations. Many favor trademark holders. For example, courts have found the likelihood of confusion requirement met even where con-

---

255. A rigorous application of the doctrine could call much of trademark law into doubt. *See* Tushnet, *supra* note 11, at 755 ("Taking modern First Amendment doctrine seriously would have significant effects on the Lanham Act, affecting everything from the standard of proof to the definition of what counts as misleading."). Thus far, however, courts have not been so inclined. *See id.* at 747 (observing that courts follow "cursory" First Amendment analysis with respect to trademark claims).

256. *See* Mattel, Inc. v. MCA Records, Inc., 296 F.3d 894, 905 (9th Cir. 2002) ("[A] trademark injunction, even a very broad one, is premised on the need to prevent consumer confusion. This consumer protection rationale—averting what is essentially a fraud on the consuming public—is wholly consistent with the theory of the First Amendment, which does not protect commercial fraud." (citing Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 566 (1980))); White v. Samsung Elecs. Am., Inc., 971 F.2d 1395, 1401 & n.3 (9th Cir. 1992) (rejecting parody defense in a case dealing with the Lanham Act and publicity claims, in part because defendant's speech was commercial); *cf.* S.F. Arts & Athletics, Inc. v. U.S. Olympic Comm., 483 U.S. 522 (1987) (upholding statute granting the United States Olympic Committee exclusive use of the word "Olympic" against a First Amendment challenge). Judge Leval has criticized reliance on the First Amendment in such cases, observing the risks of constitutionalizing more litigation than necessary. Leval, *supra* note 232, at 209.

sumers are more likely than not to avoid confusion.[257] But the flexibility of the standard also allows courts to overlook examples of *actual* confusion where appropriate.[258] The freedom to do so stems from the same vague text that allows expansive liability.[259] This method can also be used to craft doctrines that might limit trademark's reach as the Ninth Circuit's development of the nominative fair use "defense" demonstrates.

C. The Curious Case of Nominative Fair Use

The nominative fair use doctrine highlights the ambiguous status of new trademark defenses. The Ninth Circuit's creation of the doctrine demonstrates the potential for judicial creativity even when courts are constrained in developing new infringement defenses. The Third Circuit's reinterpretation of nominative fair use indicates that the implications of the Supreme Court's "trademark formalism" have yet to be fully appreciated or internalized by the lower federal courts.

1. Development of the Nominative Fair Use Doctrine

Nominative fair use reflects the simple insight that anybody should be free to refer to goods and services by their brand names. Courts handled this impulse in a variety of ways before the Ninth Circuit translated the notion into doctrine in *New Kids on the Block v. News America Publishing, Inc.*[260]

In an opinion by Judge Kozinski, *New Kids* rejected a trademark infringement claim by the band against two newspapers that had used the

---

257. *See* 4 MCCARTHY, *supra* note 26, § 23:2 (discussing necessary levels of confusion in surveys submitted as evidence to establish likely confusion).

258. *See, e.g.*, Conopco, Inc. v. May Dep't Stores Co., 46 F.3d 1556, 1564-65 (Fed. Cir. 1994) (experience of actually confused purchaser dismissed as an "atypical and an isolated incident"); *see generally* 4 MCCARTHY, *supra* note 26, § 23:13 (describing cases that have deemed examples of actual confusion as being result of carelessness or inattention).

259. On the doctrinal level, the multifactor likelihood of confusion test provides formal recognition of the view that marks in nonadjacent markets are unlikely to cause confusion and therefore are less likely to incur liability. This is perhaps an overly generous interpretation of modern practice. The multifactor test first arose as a means of analyzing likely confusion with respect to marks in non-competing markets. *See* Polaroid Corp. v. Polarad Elecs. Corp., 287 F.2d 492, 495 (2d Cir. 1961).

260. 971 F.2d 302, 308 (9th Cir. 1992). Earlier cases include, for example, *Smith v. Chanel, Inc.*, 402 F.2d 562 (9th Cir. 1968) (protecting ability of seller of a smell-alike perfume to use competitor's name in comparative advertising), and *WCVB-TV v. Boston Athletic Ass'n*, 926 F.2d 42, 46 (1st Cir. 1991) (concluding television station's unauthorized use of Boston Marathon mark to describe coverage of the sporting event would not cause confusion especially since the words were used in their descriptive sense).

group's name to conduct phone-in polls.[261] Rather than ground its holding, as the district court did, in the First Amendment,[262] or attempt to shoehorn the facts into the classic fair use defense,[263] the court focused on the trademark cause of action itself. The panel defined defendants' uses as actions that were outside of trademark law. That is, assuming certain conditions,[264] defendants' conduct was by definition unlikely to cause confusion.

> Indeed, we may generalize a class of cases where the use of the trademark does not attempt to capitalize on consumer confusion or to appropriate the cachet of one product for a different one. Such *nominative use* of a mark-where the only word reasonably available to describe a particular thing is pressed into service-lies outside the strictures of trademark law: Because it does not implicate the source-identification function that is the purpose of trademark, it does not constitute unfair competition; such use is fair because it does not imply sponsorship or endorsement by the trademark holder. "When the mark is used in a way that does not deceive the public we see no such sanctity in the word as to prevent its being used to tell the truth."[265]

---

261. *New Kids*, 971 F.2d at 304-05. Questions in the copy included, "Who's the best on the block?" "Which of the five is your fave? Or are they a turn off?" "Now which kid is the sexiest?" and "Which of the New Kids on the Block would you most like to move next door?" *Id.*

262. The panel expressed concern about the ramifications of allowing plaintiffs to police unauthorized references to the band that bear on expressive considerations. *See id.* at 306 ("[W]e need not belabor the point that some words, phrases or symbols better convey their intended meanings than others."); *id.* at 309 ("While the New Kids have a limited property right in their name, that right does not entitle them to control their fans' use of their own money."); *id.* n.9. But while expressive considerations may have exerted a gravitational pull on the final result, the panel was careful not to ground the opinion on free expression grounds, invoking the canon of constitutional avoidance. *Id.* at 305.

263. While the court noted the existence of the classic fair use defense and its incorporation in the Lanham Act, it explained that the situation raised by use of the New Kids mark to refer to the band and its members "is not the classic fair use case." *Id.* at 308.

264. The elements are: (1) "the product or service in question must be one not readily identifiable without use of the trademark;" (2) "only so much of the mark or marks may be used as is reasonably necessary to identify the product or service;" and (3) "the user must do nothing that would, in conjunction with the mark, suggest sponsorship or endorsement by the trademark holder." *Id.* at 308.

265. *Id.* at 307-08 (quoting Prestonettes, Inc. v. Coty, 264 U.S. 359, 368 (1924)). Later precedent confirmed that in the Ninth Circuit, nominative fair use is a substitute for the usual likelihood-of-confusion analysis and not a defense. *See, e.g.*, Playboy Enters., Inc. v. Welles, 279 F.3d 796, 801 (9th Cir. 2002) ("In cases in which the defendant raises a nominative use defense, the above three-factor test should be applied instead of the test

In this manner, *New Kids* avoids the textualist objections that follow from inventing a new defense to supplement the defenses in section 33(b). Rather than add to the statute's closed text, the opinion designs a prophylactic rule under one of the law's most open provisions—likelihood of confusion—to sort infringing from non-infringing acts. It is not a true defense, but an alternative method of ascertaining whether liability exists in the first place.[266] The opinion foreshadows the Supreme Court's parallel approach in *Wal-Mart*, which establishes a similar rule to sort inherently distinctive from non-inherently distinctive trade dress. In both cases, the court creates a rule to clarify how a claimant is to establish that he is entitled to protection under the Lanham Act.[267]

The Ninth Circuit's nominative fair use test could be attacked as insufficiently protective of a defendant's interest in being able to use a plaintiff's mark for referential purposes.[268] The point for present purposes is to highlight that the court found freedom to create a potentially limiting trademark doctrine in the Lanham Act's open-ended liability provisions.

2. *The Third Circuit and the Persistence of "Common Law" Thinking*

Some courts continue to innovate outside the confines of the Lanham Act's text because they have not fully appreciated the implications of the Supreme Court's recent formalist trademark jurisprudence.[269] Judges vary

---

for likelihood of confusion set forth in *Sleekcraft*."); *see also* 4 MCCARTHY, *supra* note 26, § 23:11 ("The 'nominative fair use' analysis is no more an 'affirmative defense' than is the multi-factor test of infringement used by all of the circuits.").

266. *Cf.* Pebble Beach Co. v. Tour 18 I Ltd., 155 F.3d 526, 545 (5th Cir. 1998) ("While a claim that the use was to identify the markholder's goods or services is analogous to the statutory fair-use defense, it is in actuality a claim that the use is noninfringing and thus creates no likelihood of confusion."). It is possible, of course, that a court may find that a use meets the nominative fair use test *notwithstanding* the presence of confused consumers. This would suggest that the doctrine may function like a true defense, but courts sometimes deny ordinary trademark claims notwithstanding the presence of actual confusion. *See supra* note 258.

267. One claiming trade dress in a product design could establish that the design has achieved secondary meaning with the consuming public. Similarly, the plaintiff of a nominative fair use case may argue that the defendant *has* done something to "suggest sponsorship or endorsement by the trademark holder." *New Kids*, 971 F.2d at 308.

268. Particularly in light of cases that place the burden on the defendant to demonstrate that the test is met, notwithstanding its purpose of establishing liability. *See, e.g.*, Brother Records, Inc. v. Jardine, 318 F.3d 900, 909 n.5 (9th Cir. 2003) ("[T]he nominative fair use defense shifts to the defendant the burden of proving no likelihood of confusion.").

269. As perhaps implied by the fact that some of the opinions canvassed in Section III.C resolved circuit splits notwithstanding their unanimous resolution on textualist

on the question whether they may supplement the Lanham Act with doctrines borrowed from background trademark principles or of their own creation.[270] The argument here is not that trademark cases have become

---

grounds. *See* KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc., 543 U.S. 111, 116-17 (2004); Moseley v. V Secret Catalogue, Inc., 537 U.S. 418, 428 (2003). *But cf.* Leval, *supra* note 232, at 209 ("In dealing with new challenges in the last quarter century . . . we have often read statutes with excessive literalness.").

270. For an example of borrowing background trademark principles, see *supra* note 239 and accompanying text. An example of new creations concerns the "famous marks" doctrine, which is a seldom-invoked exception to the territoriality principle. The issue concerns the priority of trademark use based on extra-territorial activity. That is, will the owner of the WIMBLEDON mark in England prevail against a user in the United States who used the mark first in the U.S., but after the mark became famous within the United States. *See* All Eng. Lawn Tennis Club (Wimbeldon) Ltd. v. Creations Aromatiques, Inc., 220 U.S.P.Q. (BNA) 1069 (T.T.A.B. 1983). In general, courts follow a strict territoriality rule. For a use of a mark to confer priority within the United States, said use must be within the nation's borders. *See, e.g.*, ITC Ltd. v. Punchgini, Inc., 482 F.3d 135, 155 (2d Cir. 2007) (noting that the territoriality principle "is basic to American trademark law. . . . Precisely because a trademark has a separate legal existence under each country's laws, ownership of a mark in one country does not automatically confer upon the owner the exclusive right to use that mark in another country"). This territoriality principle is long-pedigreed and is reflected in section 44 of the Lanham Act, which provides for the registration of, and priority for, foreign marks on the basis of foreign registration. *See* 15 U.S.C. § 1126 (2006).

A circuit split exists on the question of what to do if a foreign mark achieves fame in the United States *before* the first user engages in a use in the United States or avails herself of the procedures in section 44. In *Grupo Gigante SA De CV v. Dallo & Co., Inc.*, the Ninth Circuit relied on policy considerations to create an exception to the territoriality principle for famous foreign marks. 391 F.3d 1088 (9th Cir. 2004). The court stated:

> An absolute territoriality rule without a famous-mark exception would promote consumer confusion and fraud. Commerce crosses borders. In this nation of immigrants, so do people. Trademark is, at its core, about protecting against consumer confusion and "palming off." There can be no justification for using trademark law to fool immigrants into thinking that they are buying from the store they liked back home.

Id. at 1094 (footnote omitted).

Perhaps no justification exists, but neither does a statutory basis for the Ninth Circuit's approach, as the Second Circuit recognized in *ITC*. In confronting the same issue, *ITC* acknowledges that earlier decisions from the Trademark Trial and Appeal Board had recognized the famous marks doctrine. There is, however, "a significant concern: nowhere . . . does the Trademark Board state that its recognition of the famous marks doctrine derives from any provision of the Lanham Act or other federal law." *ITC*, 482 F.3d at 159. In light of its conclusion that neither the Act nor treaty supported the doctrine's existence, and notwithstanding the "persuasive policy argument" in support of the doctrine, *ITC* rejects it on purely formalist terms. "The fact that a doctrine may promote sound policy, however, is not a sufficient ground for its judicial recognition, particularly in an area regulated by statute." *Id.* at 165. The existence of a statute was especially relevant here "[i]n light of the comprehensive and frequently modified federal statutory

inexorably formalist, but rather that flagrant departures from the Lanham Act's text and structure are increasingly less tenable. In the Roadrunner cartoons, Wile E. Coyote remains suspended in air for a good amount of time after he runs off the cliff. At some point, however, he must look down, and that is when he inevitably falls.

Rather than looking down, the Third Circuit's approach to nominative fair use demonstrates the persistence of common law habits. In *Century 21 Real Estate Corp. v. Lendingtree, Inc.*, the court rejected the Ninth Circuit's view that nominative fair use is not a "true" defense and held that the doctrine may excuse liability despite the presence of a likelihood of confusion.[271]

The source of the holding is less clear. Judge Rendell never explained the panel's authority to announce a new defense. It therefore faces a legitimacy objection that the Ninth Circuit avoids. Because *New Kids* treats nominative use as an act that is not likely to confuse, its version of the doctrine does not need an independent grounding in the Lanham Act; the likelihood-of-confusion standard itself is the textual basis. Under this view, the frustrated trademark plaintiff has little to complain about if the test is met. If there is no likely confusion in the first instance, then she did not make her case and has no Lanham Act claim.

Not so with the Third Circuit's true "defense." The nominative defense

---

scheme for trademark protection set forth in the Lanham Act." *Id.*

271. Century 21 Real Estate Corp. v. Lendingtree, Inc., 425 F.3d 211, 221 (3d Cir. 2005) (analogizing to *KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc.*, 543 U.S. 111 (2004)); *id.* at 223 n.3 ("A nominative use defendant need only prove fairness and is not required to negate confusion."). The court made a number of modifications to the nominative test. First, it held that a court should first find whether a likelihood of confusion exists, using a modified version of the Third Circuit's multifactor test. *Id.* at 224-26. If a plaintiff makes his case, the defendant may still prevail under a modified version of the Ninth Circuit's test. Under the Third Circuit's version, the three prongs are:

> 1. Is the use of plaintiff's mark necessary to describe (1) plaintiff's product or service and (2) defendant's product or service?
>
> 2. Is only so much of the plaintiff's mark used as is necessary to describe plaintiff's products or services?
>
> 3. Does the defendant's conduct or language reflect the true and accurate relationship between plaintiff and defendant's products or services?

Id. at 228.

There is much to criticize in the Third Circuit's approach, independent of the question of whether it had the authority to create the test it did. *See, e.g.*, *id.* at 232 (Fisher, J., concurring in part and dissenting in part) (criticizing modified multifactor test because "to the extent the majority places any burden on plaintiffs at all, it is so watered-down that plaintiffs might prove likely confusion on one *Lapp* factor alone").

comes into play only if a likelihood of confusion exists.[272] But if the plaintiff proves her case, then the court cannot very well deprive her of her victory without a reason. As argued in Section IV.A.2, if a statute says "X creates liability," and a judge holds "X exists, nevertheless, there is no liability," the judge is negating the statute unless he acts pursuant to some legal authority. Regardless of whether the statute incorporates background common law defenses, there is no "gap" to be filled. The existence of the cause of action has answered the question.

Century 21 never explains what independent legal reason allows the court to negate a statutory cause of action. Instead, it focuses on the Supreme Court's conclusion in *KP Permanent* that fair use may coexist with a likelihood of confusion.[273] True enough, but *KP Permanent* addresses a statutory defense. Invoking that ruling ignores the question of whether a nominative fair use defense, in contrast to classic fair use, has any independent legal basis. One searches the opinion in vain for any such foundation, textual or otherwise.[274]

Whatever the acceptability of the Third Circuit's "common law" approach in the early decades of the Lanham Act, it is increasingly anachronistic today. If the statute does not create a general common law of unfair competition, it must be true with respect to the creation of defenses as much as to the expansion of liability. Under this view, and practical merits aside, only the Ninth Circuit's approach to nominative fair use appears legitimate against the backdrop of the Supreme Court's interpretation of the post-amendment Lanham Act.

D. Summary

Efforts to create new trademark defenses lack a stable foundation. The full consequences of this observation have yet to be internalized by the lower courts. Today, it is possible for the author of *Century 21* to agree

---

272. *Id.* at 222 ("Once plaintiff has met its burden of proving that confusion is likely, the burden then shifts to defendant to show that its use of plaintiff's mark is nonetheless fair.").

273. *Century 21*, 425 F.3d at 222-23.

274. At one point the majority suggests that it may view nominative fair use as a species of classic fair use, and thus potentially grounded in section 33(b)(4), but later indicates that it views the defenses as distinct. *Compare id.* at 221 ("Since the defendant ultimately uses the plaintiff's mark in a nominative case in order to describe its own product or services, even an accurate nominative use could potentially confuse consumers about the plaintiff's endorsement or sponsorship of the defendant's products or services.") (citation omitted), *with id.* at 222 ("Yet, the Supreme Court clearly views fair use (albeit classic fair use) as an affirmative defense.").

that "we are all textualists now"[275] and still invent a trademark defense in a decidedly nontextual opinion. But barring a shift in the Supreme Court's jurisprudence, the writing is on the wall. Sooner or later, the impact of the Supreme Court's trademark formalism and the general tenor of the age will be inescapable, leaving little room for defensive innovations that cannot be tied to actual provisions of the Lanham Act.

V. THE FUTURE OF TRADEMARK DEFENSES

This Part outlines areas of potential development for trademark defenses in light of the foregoing analysis.

A. Lanham Act Amendments

Congressional action is one obvious solution to the problems of the Lanham Act. Congress could add to the defenses available under section 33(b) of the Lanham Act or enact specific safe harbors for activities that are unlikely to cause confusion or those that may cause confusion, but whose social utility is high enough that the benefits of immunizing the acts outweigh any costs. Congress has taken this approach in its dilution legislation.[276]

One problem with a piecemeal legislative approach is that it does little to solve the fundamental problem of trademark defenses under the current act. Narrow carveouts and safe harbors will do little to counter future expansions of trademark liability. Unless the carveouts are broad,[277] the Lanham Act's underlying problem of open-ended liability provisions and narrow defenses will continue to cause difficulties in future novel contexts.

---

275. *See supra* note 2.

276. 15 U.S.C. § 1125(c)(3) (2006) provides:

The following shall not be actionable as dilution by blurring or dilution by tarnishment under this subsection:

> (A) Any fair use, including a nominative or descriptive fair use, or facilitation of such fair use, of a famous mark by another person other than as a designation of source for the person's own goods or services, including use in connection with—
>
> (i) advertising or promotion that permits consumers to compare goods or services; or
>
> (ii) identifying and parodying, criticizing, or commenting upon the famous mark owner or the goods or services of the famous mark owner.
>
> (B) All forms of news reporting and news commentary.
>
> (C) Any noncommercial use of a mark.

277. *See, e.g.*, Mattel, Inc. v. MCA Records, Inc., 296 F.3d 894, 905 (9th Cir. 2002) (construing dilution statute's defense for "non-commercial" uses).

Congress could also give judges explicit authority to devise and apply flexible defenses to trademark liability when circumstances warrant. In much the same way Congress amended the Copyright Act to incorporate the fair use doctrine, previously a common law creation,[278] it could legislate a similarly open-ended standard for judges to apply in the trademark realm. Alternatively, Congress could follow the model of the FTC Act and create a standard that contains some guidance for courts to follow in determining whether to excuse purportedly infringing conduct.[279]

B. Lanham Act Contextual "Defenses"

While the Lanham Act lacks a basis for the wholesale invention of new defenses, there is interpretive room for de facto defenses in the law's liability provisions. This Section lays a preliminary case for recognition of an explicit materiality requirement within the likelihood of confusion requirement. Consistent with the discussion in the previous Part, however, the "defenses" discussed here are not true defenses, but rather glosses on the likelihood of confusion standard. This proves to be both virtue and vice.

1. Materiality

One promising area of doctrinal development lies in giving a more overtly qualitative interpretation to the likelihood of confusion requirement. Courts could require that any alleged confusion be material before it is actionable. That is, to establish a likelihood of confusion, a trademark plaintiff must also prove that the confusion is relevant to the consuming public in making purchasing decisions.

Such a move would not be entirely novel. Materiality considerations apply to several provisions of the Lanham Act.[280] Most notably, judges have long imposed a similar materiality requirement for false advertising claims under both the current and pre-1988 versions of section 43(a).[281]

Even without an explicit requirement, materiality considerations are

---

278. *See supra* note 78.

279. The Federal Trade Commission Act authorizes the FTC to police unfair competition but provides that they may not declare a commercial practice unlawful "unless the act or practice causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition." Federal Trade Commission Act, 15 U.S.C. § 45(n) (2006). In the trademark realm, Congress could apply a similar standard or call on courts to balance the costs and benefits to consumers in the aggregate for any given class of challenged activities. *See* Grynberg, *supra* note 10, at 113-14.

280. *See supra* note 69.

281. 5 MCCARTHY, *supra* note 26, § 27:25, :35.

difficult to avoid in practice. Someone somewhere is always going to be confused about something. That fact of life plus the malleability of the likelihood of confusion standard means that a wide range of activity could trigger trademark liability. Courts must make judgment calls,[282] such as determining when confusion is *de minimis* and non-actionable.[283] Similarly, it is standard practice to assess likely confusion with the target audience in mind. We don't worry about the views of soda drinkers when determining whether a trademark for jet engines infringes.[284] We worry about whether "reasonably prudent purchasers exercising ordinary care" would be confused[285] in part because theirs is the confusion that has a marketplace impact.

Expanding these precursors into an explicit materiality requirement does not suffer from a legitimacy objection. The Lanham Act does not define "likelihood of confusion." Just as courts have always had to make quantitative assessments about what level of potential confusion amounts to "likelihood," they cannot avoid qualitative interpretations of "confusion."[286] What degree of mistaken awareness suffices for confusion? Is it conscious confusion? Subconscious? Must it be confusion that the consumer would confront while shopping, or can it be hypothesized and demonstrated through laboratory testing or with surveys?

The Lanham Act likewise does not define "origin, sponsorship, or approval," so courts must interpret those terms and their interaction with the confusion requirement.[287] While confusion as to origin or sponsorship has obvious relevance to consumers, the importance of approval is less clear depending on the precise meaning given to the term. "Confusion" as to "approval" could mean a mistaken belief that permission was required before a logo could be used on a piece of clothing apparel. Or it could be

---

282. *See* Grynberg, *supra* note 10, at 113.

283. *See supra* note 258 and accompanying text.

284. The multifactor tests of the various circuits generally consider consumer sophistication. *See* 4 MCCARTHY, *supra* note 26, §§ 24:30–:43 (listing factors used by various circuits).

285. Attrezzi, LLC v. Maytag Corp., 436 F.3d 32, 38 (1st Cir. 2006) (quoting Int'l Ass'n of Machinists & Aerospace Workers AFL-CIO v. Winship Green Nursing Ctr., 103 F.3d 196, 201 (1st Cir. 1996)); *see also* 4 MCCARTHY, *supra* note 26, § 23:91 n.1 (collecting examples of standards).

286. For example, in recognizing claims resting on initial interest or post-sale confusion, courts have justified themselves by explaining why such confusion might have a market impact. Brookfield Commc'ns, Inc. v. W. Coast Entm't Corp., 174 F.3d 1036, 1064 (9th Cir. 1999) (initial interest confusion); Ferrari S.P.A. Esercizio Fabriche Automobili E Corse v. Roberts, 944 F.2d 1235, 1244 (6th Cir. 1991) (post-sale confusion).

287. 15 U.S.C. § 1125 (2006).

more restrictive, and refer to those cases in which the markholder has placed her reputation behind the product. In choosing between the two, courts should remember that the benefits of policing consumer confusion often comes at a cost to non-confused consumers.[288] Weighing these costs and benefits may favor the more modest reading of "approval," which is permitted by the statutory text.

Finally, leaving aside the potential public policy benefits of reading a materiality requirement into the Lanham Act's open text,[289] the statute's text and context support such recognition. Textualist canons of construction support reading "approval" as something narrower than mere "permission."[290] At a broader level, trademark law's traditional consumer-protection focus militates in favor of reading the Act to actually protect consumers.[291] At the broadest level, the Supreme Court's interpretation of Article III suggests a materiality requirement. Applying the Lanham Act to activities that do not affect consumer purchases (and, by implication, sales by the trademark holder) raises a potential Article III standing issue.[292] Without material confusion the trademark plaintiff's claim of an injury in fact looks dubious. Notwithstanding the poor track record of free speech challenges to trademark law, one could make a similar claim with respect to the First Amendment. If one finds confusion in situations far removed from traditional conceptions of consumer harm, one may ask if

---

288. *See supra* note 74 and accompanying text.

289. If a goal of trademark law is to "protect the public so it may be confident that, in purchasing a product . . . , it will get the product which it asks for and wants to get," S. REP. NO. 79-1333, at 3 (1946), *as reprinted in* 1946 U.S.C.C.A.N. 1274, 1274-5, then a materiality requirement preserves that goal while creating breathing space for activities that may cause marginal confusion, but benefit a different subset of the consuming public, *see* Grynberg, *supra* note 10, at 113-14.

290. One could take an *ejusdem generis* approach and argue that the term "approval" is a general one and should be interpreted consistently with the more specific terms "origin" and "sponsorship." The same basic claim may be made under the *noscitur a sociis* canon. *See supra* note 157.

291. Even if the consumer protection goal shares time with seller protection. S. REP. NO. 79-1333, at 3, *as reprinted in* 1946 U.S.C.C.A.N. 1274, 1274-5.

292. The Court's familiar standing inquiry provides:

> [I]n order to have Article III standing, a plaintiff must adequately establish: (1) an injury in fact (i.e., a "concrete and particularized" invasion of a "legally protected interest"); (2) causation (i.e., a " 'fairly . . . trace[able]' " connection between the alleged injury in fact and the alleged conduct of the defendant); and (3) redressability (i.e., it is " 'likely' " and not "merely 'speculative' " that the plaintiff's injury will be remedied by the relief plaintiff seeks in bringing suit).

Sprint Commc'ns Co. v. APCC Servs., Inc., 128 S. Ct. 2531, 2535 (2008) (citing Lujan v. Defenders of Wildlife, 504 U.S. 555, 560-61 (1992)).

trademark liability is compatible with even the decreased First Amendment scrutiny given to commercial speech under the *Central Hudson* test.[293]

2. *Safe Harbors*

The reformist goal of increasing the number of safe harbors in trademark law may be accomplished through interpretation of the likelihood of confusion requirement.[294] The nominative fair use doctrine, which identifies a class of activities as per se unlikely to cause consumer confusion, is the model example.

Nominative fair use operates at a high level of generality, but courts could make narrower assessments of specific activities as being similarly unlikely to cause confusion. For example, trademark plaintiffs have sued Google and its clients for the search engine's practice of selling keyword advertising that enables purchasers to have their advertising returned in response to a search on a trademarked term.[295] These suits are traceable to the infamous *Brookfield* opinion, which concluded that arranging to have one's website displayed in response to a search for a trademarked term constitutes actionable "diversion" due to a misappropriation of the trademark holder's goodwill.[296] The contention that the activity satisfies the confusion requirement is largely definitional—one that may be just as easily resolved to the contrary.[297] There is room for courts to establish, if not an absolute safe harbor, a presumption that certain activities are not confusing, much like trademark law has treated comparative advertising.[298]

One may object that under the terms of the above analysis, the resulting doctrine would be less a safe harbor than a rebuttable presumption of no confusion, thus robbing the safe harbor of much of its protective force. This objection is regrettably correct and reflects a problem inherent to the

---

293. Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 566 (1980) (asking whether a regulation of commercial speech directly advances a substantial government interest). *But see supra* notes 255-256.

294. *See supra* note 13.

295. *See, e.g.*, Google Inc. v. Am. Blind & Wallpaper, No. C 03-5340 JF (RS), 2007 WL 1159950 (N.D. Cal. Apr. 18, 2007) (denying summary judgment to Google in part).

296. Brookfield Commc'ns, Inc. v. W. Coast Entm't Corp., 174 F.3d 1036, 1062 (9th Cir. 1999); *see also* Playboy Enters., Inc. v. Welles, 279 F.3d 796 (9th Cir. 2002); Soilworks, LLC v. Midwest Indus. Supply, 575 F. Supp. 2d 1118 (D. Ariz. 2008).

297. *Brookfield*, 174 F.3d at 1062 (equating diversion of consumers with improper appropriation of trademark holder's goodwill).

298. *See, e.g.*, Smith v. Chanel, Inc., 402 F.2d 562 (9th Cir. 1968). As noted above, however, sometimes courts move in the opposite direction lest they be seen to be creating defenses to trademark liability, as the Third Circuit's treatment of color in the sugar substitute market indicates. *See supra* note 132.

de facto defenses discussed in this Section.

3. *The Problem with Contextual Defenses*

A lively debate exists in the literature and the courts as to whether trademark law contains a distinct requirement that a plaintiff establish that her trademark was "used as a mark" by the junior user before proving a likelihood of confusion.[299] Opponents have a strong textual rebuttal. The classic fair use defense codified by section 33(b)(4) specifically refers to use "otherwise than as a mark."[300] But if any non-trademark use is already immune from liability, then the statutory fair use defense is superfluous.[301]

Writing in the shadow of these difficulties, Mark McKenna offers a resolution consistent with the approach described in the previous sections. He locates a trademark use requirement within the likelihood of confusion standard and argues that the Lanham Act only polices confusion as to "source," broadly defined.[302] He then defines trademark uses as those that, as a conceptual matter, may cause source confusion. "What types of uses of a trademark have the capacity to cause confusion about the source of a product or service?"[303] His answer is that "it is difficult to imagine how any use of a mark that does not indicate source could confuse consumers about source. What would cause the confusion, if not a source indication?"[304]

If Professor McKenna is correct, then the trademark use requirement has limited bite, for reasons that he himself identifies. The requirement is so inextricably bound with the underlying factual inquiry on liability as to be almost meaningless. The malleability of consumer perceptions is the very force that has helped spur the growth of trademark's scope in the past, and "source indication, like virtually everything else in trademark law, can only be determined from the perspective of consumers."[305] Basing efforts to check trademark's expansion in these same perceptions re-

---

299. *See supra* notes 8-14.

300. 15 U.S.C. § 1115(b)(4) (2006).

301. As opponents have noted. *See* Dinwoodie & Janis, *Contextualism*, *supra* note 14, at 1617. Similar textual problems arise from treating the liability requirement that a plaintiff demonstrate a "use in commerce" as being something more than a jurisdictional provision, given that the infringement provisions appear to be more expansive. *See id.* at 1609-16.

302. McKenna, *supra* note 13 (manuscript at 38-41). Professor McKenna defines "source" confusion under modern trademark doctrine to include sponsorship and affiliation relationships. *Id.* (manuscript at 39).

303. *Id.* (manuscript at 41).

304. *Id.*

305. *Id.* (manuscript at 83).

creates the underlying problem of flexible trademark liability without solving it.[306]

This difficulty, unfortunately, is common to the contextual defenses described above. Because they rely on text in the Lanham Act, they are free from the legitimacy objection that follows any wholesale invention of a defense. At the same time, they replicate the fundamental shortcoming of many existing defensive doctrines because they are bound in the vague likelihood of confusion inquiry.[307] Because they are not true defenses that would apply even if a likelihood of confusion were established, they leave room for a markholder to assert a confusion claim within the parameters of the defensive doctrine. Even if the plaintiff is ultimately unsuccessful, courts will not easily be able to dispose of cases at an early stage of litigation, eliminating much of the benefit of defensive innovations.[308]

A materiality requirement, for example, may be unable to counteract broad merchandising claims. The owner of a sports team's logo might avoid summary judgment by alleging that some consumers care whether their purchase of branded merchandise benefits the mark owner.[309] The claim may be supported by a survey that uncovers respondents who express a willingness to pay more for paraphernalia that supports the local team.[310] Even if the ultimate claim fails, it may well survive a motion to dismiss or for summary judgment. If much of trademark's current strength

---

306. *Id.* (manuscript at 82-83); *see id.* (manuscript at 64) ("[B]ecause trademark use can be determined only from the perspective of consumers, it cannot serve as a threshold requirement separable from the likelihood of confusion inquiry.").

307. *See supra* note 79.

308. *See* McGeveran, *supra* note 13, at 112-13.

309. *See* McKenna, *supra* note 13 (manuscript at 83) ("It may be that, at least in some cases, consumers do care about more than the actual source of a product or service such that other types of relationships ('sponsorship' or 'affiliation' relationships) might affect those consumers' purchasing decisions."); Mark P. McKenna, *The Normative Foundations of Trademark Law*, 82 NOTRE DAME L. REV. 1839, 1916 (2007) ("Producers are able to frame just about any argument for broader protection in terms of consumer expectations, which they are in position to influence systematically through marketing.").

310. The holder of the Michelob beer mark used survey evidence to a similar effect in *Anheuser-Busch, Inc. v. Balducci Publns.*, 28 F.3d 769 (8th Cir. 1994). The case involved a humor magazine's satirical use in a fake advertisement of the beer marks as being drenched in oil. The court concluded a likelihood of confusion existed, relying in large part on survey evidence indicating that most viewers of the ad thought that mark holder's permission was required. *Id.* at 772-73. In addition, "[f]ifty-five percent construed the parody as suggesting that Michelob beer is or was in some way contaminated with oil. As a result, twenty-two percent stated they were less likely to buy Michelob beer in the future." *Id.* at 773. A future plaintiff could try to build a case for material confusion by asking instead if a prospective beer buyer would be less likely to purchase the plaintiff's product if the plaintiff had licensed her mark for an unappealing use.

comes from the *in terrorem* threat of litigation that cannot be easily turned away at an early stage, then a materiality requirement grounded in the likelihood of confusion requirement may prove less useful than hoped.[311]

C. Roll Back Trademark's Expansion

If courts are incapable of devising new "true" defenses to trademark infringement, and quasi-defenses tied to the Lanham Act's liability-creating provisions are inherently weak, what's left? The problem with which we began: trademark's expansion. Rather than attempting to devise defensive doctrines to cabin trademark's expanded reach, reformist efforts may at times be better spent on the expansion itself.

Hope on this front stems from more than the various academic proposals for reforming trademark law. Courts might take their cue from the Supreme Court's recent reluctance to entertain expansive liability claims under the Lanham Act.[312] Many especially expansive rulings have received judicial as well as academic criticism, and removal of their deleterious consequences could just be an *en banc* review away.[313]

Optimism is tempered by the knowledge that any change in judicial policy must accommodate the text of the Lanham Act, which reflects trademark's past gains. So while the Supreme Court suggested, and eventually held, that color could not be trademarked without first establishing secondary meaning,[314] it viewed itself bound to the proposition that color can be trademarked in the first place.[315] Similarly, courts may give a re-

---

311. *See supra* note 73 and accompanying text.

312. *See supra* Section III.B. And indeed it may be argued that the current trend is in favor of defendants when expressive uses of trademarks are at issue. *See* McGeveran, *supra* note 13, at 61.

313. Playboy Enters., Inc. v. Netscape Commc'ns Corp., 354 F.3d 1020, 1034-35 (9th Cir. 2004) (Berzon, J., concurring). Judge Berzon expressed "concern that [*Brookfield*] was wrongly decided and may one day, if not now, need to be reconsidered *en banc*." She argued:

> There is a big difference between hijacking a customer to another website by making the customer think he or she is visiting the trademark holder's website (even if only briefly), . . . and just distracting a potential customer with another choice, when it is clear that it is a choice.

Id.

314. *See* Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 163 (1995) ("We cannot find in the basic objectives of trademark law any obvious theoretical objection to the use of color alone as a trademark, where that color has attained 'secondary meaning' and therefore identifies and distinguishes a particular brand (and thus indicates its 'source').."); *see also* Wal-Mart Stores, Inc. v. Samara Bros., 529 U.S. 205, 212 (2000) ("We held that a color could be protected as a trademark, but only upon a showing of secondary meaning." (citing *Qualitex*, 514 U.S. at 162-63)).

315. *Qualitex*, 514 U.S. at 162 ("Since human beings might use as a 'symbol' or 'de-

strictive reading to the word "approval," but they cannot read the term out of the statute.[316] Finally, to the extent that the structure of trademark litigation stacks the deck in favor of trademark holders, and by extension expansive trademark doctrines,[317] these tendencies will likely prove resistant to anything but a fundamental shift in judicial attitudes toward expansive trademark claims.

VI. CONCLUSION

Efforts to reign in trademark law through the creation of defensive doctrines need open text to survive. Such text is found in the Lanham Act's liability-creating clauses, particularly the likelihood of confusion standard. While such provisions do indeed have potential to help create de facto defenses, reliance upon them threatens to replicate many of the defects of current trademark jurisprudence. In particular, it is difficult to rely on the malleable likelihood of confusion standard to create a defense without importing the problems of indeterminacy that already complicate the swift resolution of trademark cases.

Absent congressional action, courts have only a limited ability to correct the imbalance between trademark liability and defenses. The alternative is to roll back the expansions of liability that have already occurred. But this approach also faces the tilted playing field of the Lanham Act's textual defense/liability mismatch. Even then, it is not just a matter of convincing judges to "switch sides" with respect to views of expansive trademark. There remains the matter of the accumulated precedent to date. Any judicially paced effort may therefore be little more than a rearguard action. Things are worse than we think.

---

vice' almost anything at all that is capable of carrying meaning, this language, read literally, is not restrictive.") (discussing 15 U.S.C. § 1127).

316. *See supra* Section V.B.1.

317. Grynberg, *supra* note 10, at 64-87 (arguing that the structure of trademark litigation drives trademark's expansion).

DETHRONING *LEAR?* INCENTIVES TO INNOVATE AFTER *MEDIMMUNE*

By Rochelle Cooper Dreyfuss[†] *and Lawrence S. Pope*[‡]

TABLE OF CONTENTS

In January 2007, the U.S. Supreme Court radically changed the rules on licensing U.S. patents to the substantial disadvantage of anyone hoping to derive reliable revenues from the extraction of royalties. At the time of the decision in *MedImmune v. Genentech*,[1] the law was fairly well-established: a licensee in good standing—that is, a licensee who was not in breach of its licensing agreement and therefore not at risk of having the license terminated—was barred from challenging the validity of the licensed patent.[2] To gain access to federal court, such a licensee had to

---

† Pauline Newman Professor of Law, New York University School of Law. The author wishes to thank Michael Risch for his thoughtful comments, Andrew Michaels, NYU Law School, Class of 2010 for research assistance, and the Filomen D'Agostino and Max E. Greenberg Research Fund for financial support.

‡ A solo practitioner. Comments welcome at lawrences2005@yahoo.com.

1. MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118 (2007).

2. *See, e.g.*, Gen-Probe Inc. v. Vysis, Inc., 359 F.3d 1376 (Fed. Cir. 2004); Cordis Corp. v. Medtronic, Inc., 780 F.2d 991 (Fed. Cir. 1985).

create a "case" or "controversy"[3] by committing a breach of the agreement, usually by withholding royalties. In *MedImmune*, however, the Supreme Court opened the federal courthouse doors to licensees in good standing by allowing them to bring declaratory judgment actions challenging patent validity.

In many ways, the decision in *MedImmune v. Genentech* is unsurprising. It is consistent with the Court's view on justiciability more generally and with the availability of declaratory relief in other legal contexts.[4] The decision also resolves tensions with other procedural rules for patent litigation.[5] For example, it has long been clear that a licensee in good standing can bring a declaratory judgment action to resolve a dispute on whether a given product or activity falls within the scope of the claims of the licensed patent.[6] Since disputes about patent scope often trigger questions of validity, the distinction drawn in the pre-*MedImmune* era was always somewhat illusory.

As important, the Supreme Court recently expressed substantial concern about the strength and prevalence of patents. In the last three Terms, it raised the standard of inventiveness,[7] limited the patentee's right to injunctive relief,[8] circumscribed the geographic reach of infringement liability,[9] questioned the range of patentable subject matter,[10] and expanded the scope of both the exhaustion doctrine[11] and the statutory experimental-use defense.[12] By arguably opening the federal courts to patent challenges by

---

3. U.S. CONST. art. III; Declaratory Judgment Act, 28 U.S.C. § 2201(a) (2006).

4. *See* Sprint Commc'ns Co. v. APCC Servs., Inc., 128 S. Ct. 2531 (2008); *MedImmune*, 549 U.S. at 126–27, 137 n.11 (discussing Aetna Life Ins. Co. v. Haworth, 300 U.S. 227 (1937)); *see also* Teva Pharm. USA, Inc. v. Novartis Pharm. Corp., 482 F.3d 1330, 1346–47 (Fed. Cir. 2007) (Friedman, J., concurring).

5. The *MedImmune* Court specifically mentioned *Cardinal Chem. Co. v. Morton Int'l, Inc.*, 508 U.S. 83 (1993), which held that a decision of noninfringement, even if it resolves a case, nonetheless does not moot a declaratory judgment counterclaim of patent invalidity. *See MedImmune*, 549 U.S. at 133 n.11.

6. *See, e.g.*, Westinghouse Elec. & Mfg. Co. v. Formica Insulation Co., 266 U.S. 342, 351 (1924).

7. KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398 (2007).

8. eBay Inc. v. MercExchange, L.L.C., 547 U.S. 388 (2006).

9. Microsoft Corp. v. AT&T Corp., 550 U.S. 437 (2007).

10. Lab. Corp. of Am. Holdings v. Metabolite Labs., Inc., 370 F. 3d 1354 (Fed. Cir. 2004), *cert. denied*, 548 U.S. 124 (2006) (Breyer, J., dissenting).

11. Quanta Computer, Inc. v. LG Elecs., Inc., 128 S. Ct. 2109 (2008).

12. Merck KGaA v. Integra Lifesciences I, Ltd., 545 U.S. 193 (2005). To be sure, patentees have also won a few cases: in *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 535 U.S. 722 (2002), the Supreme Court reversed the Federal Circuit's constrained view of the doctrine of equivalents, and in *Illinois Tool Works Inc. v. Indepen-*

licensees—indeed, by anyone with a substantial investment in the technology covered by the patent[13]—*MedImmune* appears to create new avenues for courts to police the implementation of patent law and to release advances that should not have been patented into the public domain. As such, it extends the rationale of *Lear v. Adkins*,[14] a forty-year-old case in which the Supreme Court rejected the contract-law doctrine of licensee estoppel and allowed a licensee that had repudiated its agreement to challenge the patent. According to the *Lear* Court:

> Licensees may often be the only individuals with enough economic incentive to challenge the patentability of an inventor's discovery. If they are muzzled, the public may continually be required to pay tribute to would-be monopolists without need or justification. We think it plain that the technical requirements of contract doctrine must give way before the demands of the public interest . . . .[15]

Still, the *MedImmune* decision is, in significant ways, astonishing. As Sean O'Connor pointed out, the reasoning is ahistorical, convoluted, and blind to the realities of licensing practice.[16] It effects a dramatic change in the rules of the licensing game by substantially enhancing the bargaining position of the licensee to the detriment of the patent holder. The licensee can now seek a new arrangement any time it can mount a credible contract dispute. Furthermore, it can do so without taking any real risk, for if the patent is upheld, the licensee can continue to rely on the license. At the same time, the patent holder is trapped in a difficult situation. It is tied to a deal that is unraveling and encumbered with substantial risk: because the Supreme Court also abolished the doctrine of mutuality of estoppel,[17] any

---

dent Ink, Inc., 547 U.S. 28 (2006), the Court made it more difficult to prove that the patentee engaged in illegal tying.

13. *See, e.g.*, SanDisk Corp. v. STMicroelectronics, Inc., 480 F.3d 1372 (Fed. Cir. 2007) (holding that conduct related to utilization of the patented invention can create a justiciable controversy, even when the patent holder denies a plan to sue). The Federal Circuit may, however, be retreating from this broad view. *See* Janssen Pharmaceutica, N.V. v. Apotex, Inc., 540 F.3d 1353, 1362–63 (Fed. Cir. 2008) (requiring more than speculative fear of harm to establish that the dispute is "definite and concrete"); Prasco, L.L.C. v. Medicis Pharm. Corp., 537 F.3d 1329 (Fed. Cir. 2008) (requiring the plaintiff to show an affirmative act by the patentee that demonstrates an intent to sue).

14. Lear, Inc. v. Adkins, 395 U.S. 653 (1969).

15. *Id.* at 670–71.

16. Sean M. O'Connor, *Using Stock and Stock Options to Minimize Patent Royalty Payment Risks After* MedImmune v. Genentech, 3 N.Y.U. J. L. & BUS. 381, 389–422, 429–34, 443–51 (2007).

17. Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313 (1971).

decision on patent invalidity will be good not only against the challenger, but also against the world.

Strangely, although the *MedImmune* decision is understood as part of the patent-law canon, there is virtually nothing in the opinion that rationalizes the result in terms of innovation policy.[18] Indeed, if the decision is taken at face value, it would be hard to articulate a strong justification for the ways in which it changes licensing practice.[19] True, it appears that the public will benefit from *MedImmune* because the decision effectively anoints a new group of "private attorneys general" with freedom to patrol the patent landscape and invalidate patents. But the asymmetries in the parties' bargaining positions will ultimately endanger the public interest in scientific progress. At the time of a challenge, the risk that the patent will be invalidated could lead the patent holder to settle on highly unfavorable terms. In such cases, the patent will remain in force. Accordingly, society will not gain free access to the invention. The patent holder will, however, lose revenue, leading to an impairment of patent value and a decrease in incentives to invent.

Furthermore, as John Schlicher has argued,[20] as patent holders begin to understand the decision, they will raise licensing fees to compensate for assuming a new risk, or, as O'Connor has suggested, they will require licensees to pay the full cost of the license upfront.[21] In either event, fewer deals will likely be made and fewer inventions will be put to socially-beneficial uses.[22] Worrisomely, these effects will be magnified in the emerging sectors (such as bioinformatics) that are the most critical to technological progress. In these areas, the value of the invention is often difficult to calculate, making running royalties the only efficient way to allocate benefits. At the same time, however, patentability law tends to be so highly unsettled that it is easy to mount credible challenges. To make

---

18. Interestingly, patent policy was extensively briefed by the parties and their amici. *See* Erik Belt & Keith Toms, *The Price of Admission: Licensee Challenges to Patents After* MedImmune v. Genentech, 51-JUN. BOSTON B.J. 10, 10, 12 (2007).

19. *MedImmune* has some implications with important public benefits. *See infra* notes 43–53 and accompanying text.

20. John W. Schlicher, *Patent Licensing, What to Do After* MedImmune v. Genentech, 89 J. PAT. & TRADEMARK OFF. SOC'Y 368 (2007).

21. O'Connor, *supra* note 16, at 452–55; *see also* William H. Hollander, *Challenging Patents Becomes Easier*, 54-APR. FED. LAW. 18, 19 (2007).

22. *See* Stephanie Chu, *Operation Restoration: How Can Patent Holders Protect Themselves From* MedImmune?, 2007 DUKE L. & TECH. L. REV. 0008, ¶¶ 28–33; Jeffrey M. Butler & Ashton J. Delauney, *The Implications of* MedImmune v. Genentech, INTELL. ASSET MGMT. (IAM) MAGAZINE, Feb. 2007 (Supplement), at 37, http://www.iam-magazine.com/issues/article.ashx?g=5206255c-cbee-4de8-8426-9ea0d951a3c9.

matters even worse, the participants are likely to be in poor positions to deal with either risk or cash-flow problems: the patent holders in emerging sectors are likely to be small companies that are highly dependent on their patent revenues, or universities, which rely on licensing income to fund their technology transfer operations. Meanwhile, the licensees could be cash-starved startups that cannot afford to pay for the license until their marketing efforts pay off.

Nor is it clear how legally-sophisticated and financially-well-endowed firms should deal with the problems generated by *MedImmune*, for the decision leaves many important issues hanging. Although the opinion appears to extend *Lear*, the Court twice noted that it was leaving various aspects of the licensee estoppel doctrine open.[23] In addition, the Court insisted that its analysis was unaffected by whether the "case" it considered justiciable was a patent dispute (over validity and infringement) or a contract dispute (over whether royalties were owed under the license).[24] Yet the Court supplied a rationale for finding the controversy justiciable only on the basis of the contract dispute—and then proceeded to leave the exact nature of that dispute murky.[25] The opinion says the contested issue was whether the contract provided for the payment of royalties even if the patent was invalidated.[26] But it is difficult to understand how that could be a litigable issue. As between the parties, the contract settled the question by explicitly providing that royalties were owed only on patents "which have neither expired nor been invalidated."[27] Besides, existing case law holds that it is misuse to demand royalties that extend beyond the life of the patent.[28] Even if it were plausible to interpret the contract as requiring post-invalidation payments, there would be a problem. *Lear* abrogated the doctrine of licensee estoppel in order to encourage validity challenges. But if, on remand, the trial court were to agree that payment was owed irrespec-

23. MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 124 (2007) ("We express no opinion on whether a *nonrepudiating licensee* is similarly relieved of its contract obligation during a successful challenge to a patent's validity—that is, on the applicability of licensee estoppel under these circumstances."). *Id.* at 128 ("Assuming (without deciding) that respondents here could not claim an anticipatory breach and repudiate the license, the continuation of royalty payments makes what would otherwise be an imminent threat at least remote, if not nonexistent.").

24. *Id.* at 123.

25. *Id.* at 123; *see also id.* at 140 ("[T]he Court never explains what the supposed contract dispute is actually about.") (Thomas, J., dissenting).

26. *See, e.g.*, MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 125 (2007).

27. *See infra* note 36 and accompanying text.

28. *See* Brulotte v. Thys Co., 379 U.S. 29, 32 (1964).

tive of whether the patent was valid, it would never reach the validity issue, and the goal of *Lear* would be frustrated.[29]

In fact, the best reading of *MedImmune* is that it reverses *Lear* or, more accurately, the received wisdom on what *Lear* means. Thus, while *Lear* is understood as prohibiting the enforcement of any contract provision that reduces the licensee's incentive to challenge validity,[30] *MedImmune* can be interpreted as permitting patent holders to bargain for such restrictions. The Court's analysis of the contract issue suggests that patent holders could avoid challenge by nonrepudiating licensees by simply omitting reference to the validity of the licensed patent.[31] Moreover, by specifically noting that "it is not clear where [in the contract] the prohibition against challenging the validity of patents is found,"[32] the Court implied that if it had found a no-contest provision, it would have enforced it.

Although this interpretation of *MedImmune* is counterintuitive in that it permits licensees to be "remuzzled," it does so only when the licensee has notice that its opportunity to challenge the patent is truncated. Not only does this interpretation have the salutary effect of spurring the licensee to fully vet the patent before it is licensed, it also comports rather well with many of the jurisprudential developments that have occurred since *Lear* was decided—with modern economic theory about the monopolizing potential of patents and the problems of risk allocation, with increased appreciation for the need to promote technology transfer, as well as with current views on the interface between federal intellectual property law and

---

29. *MedImmune*, 549 U.S. at 124 n.3. The resolution of the case also raises a difficult issue concerning subject matter jurisdiction. The Declaratory Judgment Act does not confer subject matter jurisdiction; its sole function is to change the time when a suit can be brought. Thus, the basis for jurisdiction must be found in other law. Genentech is a California corporation and MedImmune is located in Maryland; accordingly, there was diversity jurisdiction in the actual case. *See* 28 U.S.C. § 1332 (2006). Had the parties both been domiciled in the same state, the only possible bases for jurisdiction would be federal question jurisdiction, 28 U.S.C. § 1331 (2006), or patent jurisdiction, 28 U.S.C. § 1338 (2006). But if the dispute was justiciable because of a contract issue, it is not clear that the federal issue would have been prominent enough to put the case on the federal docket. *See, e.g.*, Merrell Dow Pharm. Inc. v. Thompson, 478 U.S. 804 (1986); Franchise Tax Bd. of Cal. v. Constr. Laborers Vacation Trust for So. Cal., 463 U.S. 1 (1983). *But see* Grable & Sons Metal Prods., Inc. v. Darue Eng'g & Mfg., 545 U.S. 308 (2005) (taking an interest-based approach to the issue of federal question jurisdiction).

30. *See* Rochelle Cooper Dreyfuss, *Dethroning* Lear*: Licensee Estoppel and the Incentive to Innovate*, 72 VA. L. REV. 677, 694–95 (1986) (citing post-*Lear* cases); *see also* Pope Mfg. Co. v. Gormully, 144 U.S. 224 (1892) (refusing to enforce a licensee's promise not to contest the validity of the licensed patent).

31. *See* Schlicher, *supra* note 20, at 367.

32. *MedImmune*, 549 U.S. at 135.

state contract law.[33] Policing the patent landscape remains an important objective, but as one of us pointed out over twenty years ago, relaxing the requirements for declaratory judgment relief and reforming procedures within the Patent and Trademark Office ("PTO") provide effective avenues for testing patent validity without interfering with the bargains that are crucial to the dissemination of knowledge.[34]

Part II of this Article examines the consequences of *MedImmune* and sets out the arguments for reading it as promoting private ordering, even when the arrangements impair the propensity of licensees to bring validity challenges. Part III makes the arguments about the vitality of *Lear* and its lore more concrete by discussing five general approaches to contracting in light of *MedImmune*, each representing a different view of *MedImmune*'s effect on prior case law. The Article concludes that *MedImmune* is best read as giving parties considerable freedom to tailor relationships to meet business needs. This freedom is, however, based on the notion that party autonomy usually serves an important public purpose. But the public interest is not always served by private arrangements. For example, in opening the door to more patent challenges, *MedImmune* suggests that the public interest in settling disputes does not outweigh other social goals. Thus, the decision casts doubt on a line of cases upholding agreements between branded pharmaceutical companies and generic manufacturers which eliminate competition in the name of reducing litigation costs.[35]

---

33. *See infra* notes 74–82 and accompanying text.

34. 35 U.S.C. §§ 311–312 (2006) (ex parte reexamination); 35 U.S.C. §§ 313–316 (2006) (inter partes reexamination); Dreyfuss, *supra* note 30, at 756–64 (discussing declaratory relief and reexamination). Congress also contemplated the institution of a post-grant opposition procedure which would significantly expand the scope of *inter partes* validity challenges in the PTO from that available via inter partes reexamination. *See* Patents Depend on Quality Act of 2006, H.R. 5096, 109th Cong. (2006); H.R. 5418, 109th Cong. (2006); Patent Reform Act of 2005, H.R. 2795, 109th Cong. (2005); *see also* John E. Calfee & Claude Barfield, *Congress's Patent Mistakes*, WALL ST. J., Oct. 29, 2007, at A18 (suggesting that the opposition procedure is one of the few "useful and innovative" reforms in the bills).

35. *See, e.g.*, *In re* Tamoxifen Citrate Antitrust Litig., 466 F.3d 187, 202 (2d Cir. 2006); Schering-Plough Corp. v. F.T.C., 402 F.3d 1056, 1068 (11th Cir. 2005); Valley Drug Co. v. Geneva Pharm., Inc., 344 F.3d 1294, 1304 (11th Cir. 2003) (rejecting antitrust challenges to settlement agreements); *see also* S. 369, 111th Cong. (2009) (proposing to ban brand name pharmaceutical firms from making reverse-payment settlements in order to keep generics off the market).

I. *MEDIMMUNE* AND ITS RAMIFICATIONS FOR LAW, LORE, AND LICENSING

A. The Case

Both of the parties in *MedImmune* operate in the biotechnology sector: Genentech, one of the founders of the field, generates upstream genetic information, which it uses and licenses to others for the purpose of finding new biotherapeutic products; MedImmune develops medicines to treat a variety of diseases. The parties' 1997 agreement involved two sets of advances, one covered by a patent that had issued prior to the license; the other was, at the time of the negotiations, the subject of a pending patent application. MedImmune agreed to pay royalties on all "licensed products," defined to include:

> [any product] the manufacture, use or sale of which . . . would, if not licensed under th[e] Agreement, infringe one or more claims of either or both of [the covered patents,] which have neither expired nor been invalidated by a court or other body of competent jurisdiction from which no appeal had been or may be taken.[36]

The license agreement gave MedImmune the right to terminate on six months' notice.

At the outset of the relationship, MedImmune did not utilize any of Genenetech's technology. However, four years into the agreement, a patent was awarded on the pending application. Genentech immediately sent MedImmune a letter expressing its belief that the new patent covered Synagis, a drug that prevented respiratory tract disease in infants and young children and which accounted for 80% of MedImmune's revenue. Although MedImmune thought the patent was invalid, unenforceable, and not infringed, it regarded the letter as a threat to terminate the license and sue for infringement. Fearing monetary liability (including treble damages and attorneys' fees) and loss of the bulk of its business, MedImmune paid "under protest and with reservation of all of [its] rights."[37] It then brought a declaratory judgment action to clarify its position. At the time, Federal Circuit case law required a declaratory plaintiff to establish a controversy by showing that it had a basis for fearing an imminent suit by the patentee and by demonstrating activity that constituted infringement or the intent to infringe.[38] Since MedImmune was a licensee in good standing, the district court reasoned that it was not in apprehension of a suit. Accordingly, the

36. *MedImmune*, 549 U.S. at 121.
37. *Id.* at 122.
38. *See, e.g.*, Gen-Probe Inc. v. Vysis, Inc., 359 F.3d 1376, 1379 (Fed. Cir. 2004).

court dismissed the claims on subject matter jurisdiction—in fact, justiciability—grounds; the Federal Circuit affirmed.[39]

The Supreme Court reversed. Justice Scalia spent the first part of the opinion bringing patent law into line with justiciability doctrine in other fields. He began by rehearsing the Court's test for justiciability— that the "dispute be 'definite and concrete, touching the legal relations of parties having adverse legal interests'; and that it be 'real and substantial' and 'admi[t] of specific relief.' "[40] He ended, in a footnote, by rejecting the Federal Circuit's apprehension-of-suit requirement.[41] In the main part of the opinion, the Court examined the circumstances to determine whether MedImmune's continuing royalty payments meant that the dispute was not justiciable. It reasoned that if MedImmune were required to repudiate the agreement in order to put the validity and infringement of the patent into issue, it would, in effect, be put to a choice—risk its business or abandon its rights. Since the precise goal of the Declaratory Judgment Act is to relieve the plaintiff of the obligation to "bet the farm,"[42] the Court held that even without repudiation, there was a "case of actual controversy" within the meaning of the Act and a "case" or "controversy" within the meaning of the Constitution.

In many ways, *MedImmune* is a welcome decision. If there were general confidence in patent law doctrine and in the workings of the PTO, the Federal Circuit's apprehension-of-suit requirement would supply an important limit on the availability of declaratory relief. Erecting a high barrier to suit protects patent value because it prevents those with an interest in the field of the invention from wearing down the patent holder with successive—and expensive—lawsuits. In the licensing context, a high barrier equalizes the risks of the licensee and the patent holder: the patent holder's risk of having the patent invalidated against the world is balanced by the requirement that the licensee repudiate the contract and risk the possibility that it will no longer enjoy the right to practice the patented invention. In the *MedImmune* Court's words, *both* sides must bet the farm. The result promotes settlement (or, better, respect for the agreement).

However, the current climate is one in which patents are regarded with considerable suspicion. Observers question the capacity of the PTO to conduct adequate examinations and thus view litigation as an important

39. *MedImmune*, 549 U.S. at 122.

40. *Id.* at 126–127 (citing Aetna Life Ins. Co. v. Haworth, 300 U.S. 227, 240–41 (1937)).

41. *Id.* at 132 n.11.

42. *Id.* at 129.

avenue for testing validity.[43] As noted earlier, the Supreme Court has also evinced doubts about the law the PTO is required to apply.[44] In that situation, the Federal Circuit's apprehension-of-suit test works a great deal of mischief. Suspecting that their patents may be invalid, patent holders are reluctant to actually permit them to be challenged in court. Nonetheless, the high barrier to suit allows these patents to be used to chill competition. Patent holders can send potential users ambiguous messages—"discussions" of their patent positions—that interfere with customer relationships and discourage investment in the infrastructure necessary to practice or market the invention.[45] As long as these "discussions" fall short of explicit threats to sue, the apprehension-of-suit test prevents recipients from going to court to clear their positions. Even if the patent holder allows a case to go to trial, it can have souring litigation dismissed prior to judgment by simply interposing a promise not to assert claims related to existing activity.[46]

National health policy is also endangered. The Hatch-Waxman Act[47] promotes competition in the pharmaceutical sector by giving a generic producer the right to apply for clearance to market a patented drug during

---

43. *See generally*, ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS (2004); Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495 (2001).

44. *See supra* notes 7–10 and accompanying text; *see also* Rochelle Cooper Dreyfuss, *In Search of Institutional Identity: The Federal Circuit Comes of Age*, 23 BERKELEY TECH. L.J. 787, 791 (2008) ("[B]ecause the Supreme Court has either reversed or vacated virtually all the Federal Circuit patent decisions that it has recently reviewed, 'dissatisfied' may be a more accurate description of its attitude.").

45. For an example of this type of activity, see the facts of *SanDisk Corp. v. STMicroelectronics* NV, 480 F.3d 1372, 1374–76 (Fed. Cir. 2007). *Cf.* Arrowhead Indus. Water, Inc. v. Ecolochem, Inc., 846 F.2d 731, 735 (Fed. Cir. 1988) (noting that without the possibility of declaratory relief, competitors were "victimized" by patent owners who engaged in "extra-judicial patent enforcement with scare-the-customer-and-run tactics that infect[ed] the competitive environment of the business community with uncertainty and insecurity.").

46. For an example of this type of activity, see the facts of *Benitec Austl., Ltd. v. Nucleonics, Inc.*, 495 F.3d 1340, 1342–43 (Fed. Cir. 2007). *See also* Caraco Pharm. Labs., Ltd. v. Forest Labs., Inc., 527 F.3d 1278 (Fed. Cir. 2008). Also of interest is the family of consolidated cases involving Columbia University's Richard Axel patent. After Judge Wolf commented unfavorably on the validity of the patent in *Biogen Idec MA Inc. v. Trustees of Columbia University*, 332 F. Supp. 2d 286, 297 (D. Mass. 2004), Columbia sought to end the litigation by granting covenants not to sue. *See In re* Columbia Univ. Patent Litig., 343 F. Supp. 2d 35 (D. Mass. 2004).

47. Drug Price Competition and Patent Term Restoration ("Hatch-Waxman") Act of 1984, Pub. L. No. 98-417, 98 Stat. 1585 (2006) (including scattered amendments to titles 15, 21, 28, and 35 of the United States Code).

the patent period. The Act also allows a proprietary manufacturer to treat such an application as patent infringement and sue to keep the generic off the market.[48] However, because of the requirements of the apprehension-of-suit test, rather than permit a court to test the validity of its entire patent position relative to the drug in question, a proprietary manufacturer can sue on only a few of the relevant patents. Without the ability to bring a declaratory judgment action to deal with the other patents, the generic is left without the assurance it needs to launch its (much cheaper) products.[49]

Under *MedImmune*, however, these tactics are less likely to be successful. Several cases have reached the Federal Circuit in the aftermath of *MedImmune*. In each case in which the court thought this form of gamesmanship was going on, it permitted the litigation to go to judgment.[50] At the same time, however, the court has been able to protect patent holders from abusive practices. It has relied on the standard justiciability requirements of concreteness, adversariness, and redressability to limit the range of potential declaratory plaintiffs to those who have made a substantial investment in the patented technology.[51] In addition, the court has stressed that the Declaratory Judgment Act gives trial courts discretion to dismiss cases even when they are justiciable. That discretion, the court has suggested, is properly exercised in situations where the declaratory plaintiff is behaving opportunistically.[52] The court has also noted that patent holders

---

48. 21 U.S.C. § 355 (2006); 35 U.S.C. § 271(e) (2006); *see* 21 U.S.C. § 355(j)(5)(C) (creating a "[c]ivil action to obtain patent certainty").

49. *See, e.g.*, *Caraco*, 527 F.3d at 1296–1297; Teva Pharm. USA, Inc. v. Novartis Pharm. Corp., 482 F.3d 1330, 1243 (Fed. Cir. 2007).

50. *See Caraco*, 527 F.3d at 1297; *Teva*, 482 F.3d at 1346; *SanDisk*, 480 F.3d at 1383. *But see* Prasco LLC v. Medicis Pharm. Corp., 537 F.3d 1329, 1341 (Fed. Cir. 2008) (requiring that the plaintiff show an affirmative act by the patentee demonstrating an intent to sue). Cf. *Benitec*, 495 F.3d at 1351 (Dyk J., dissenting from a dismissal of a declaratory judgment action on the ground that "patentee's manipulative efforts to defeat declaratory jurisdiction are clear enough").

51. *See* Cat Tech L.L.C. v. TubeMaster, Inc., 528 F.3d 871, 880 (Fed. Cir. 2008) (noting that until the declaratory plaintiff has taken substantial steps toward utilizing the patent, the facts are too fluid to entertain an action); *see also* Benitec, 495 F.3d at 1346 (describing that when research protected by the research exemption, 35 U.S.C. § 271(e)(1) (2006), is underway, the suit is not ripe); *cf.* Janssen Pharmaceutica, N.V. v. Apotex, Inc., 540 F.3d 1353, 1360–61 (Fed. Cir. 2008) (requiring the declaratory plaintiff to demonstrate that the alleged harm was caused by the dispute over the validity of the patent in issue); Adenta GmbH v. OrthoArm, Inc., 501 F.3d 1364, 1369–70 (Fed. Cir. 2007) (verifying that the relationship between the parties was adversarial).

52. *See, e.g.*, Sony Elecs., Inc. v. Guardian Media Techs., Ltd., 497 F.3d 1271, 1289 (Fed. Cir. 2007) (suggesting that if the facts supported the contention that there was "a strategic motive to obtain a more favorable bargaining position in the ongoing negotiations with the patentee and also to undermine the value of the patent so as to impede its

can sometimes protect themselves from a multiplicity of suits by using the transfer provisions of federal law to consolidate all related cases before a single tribunal.[53]

B. Ramifications for Licensing

Despite these advantages for the administration of patent law generally, *MedImmune* raises substantial concerns in the licensing context. As explained in the Introduction, the holding substantially alters the parties' bargaining position. In its effort to save the licensee from having to "bet the farm," the Court shifted the entire litigation risk to the patent holder. Because of *Blonder-Tongue* and the demise of mutuality of estoppel,[54] a declaration of patent invalidity is good against the world, and not just against the challenger. Accordingly, an adverse decision will destroy the entire income stream flowing from the patent. Knowing the risk, patent holders are much more likely to settle improvidently. As a result, the patent will remain in force, but the licensee will now capture part of the supracompetitive return associated with it. Since patents are rarely vulnerable on the ground that the *licensee* was the true inventor, payments wind up going to the wrong party and incentives to innovate are weakened.

It is not even clear that pre-litigation settlements with licensees will stick, for the question whether settlements should be treated like consent decrees—which are presumably entitled to res judicata effect[55]—or like any other contract has long been controversial.[56] The rationale for barring the settling licensee from attacking the patent is that the public interest in avoiding litigation is strong enough to overcome *Lear*'s goal of opening the courthouse door to validity challenges. But since the *MedImmune* Court surely knew that its decision was bound to produce considerable

---

sale or licensing to a third party," the trial court could dismiss the case; no such motive was, however, shown); *see also* Cellco P'ship v. Broadcom Corp., No. 2006-1514, 2007 WL 841615, at *1 (Fed. Cir. Mar 19, 2007) (suggesting that a court could exercise its discretion to dismiss a case when multiple related actions are pending).

53. Micron Tech., Inc. v. Mosaid Techs., Inc., 518 F.3d 897, 903–905 (Fed. Cir. 2008) (explaining that "the possibility of consolidation with related litigation" qualifies as a factor courts must consider when weighing transfer between forums under 28 U.S.C. § 1404(a)).

54. Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313, 333 (1971).

55. *See, e.g.*, Interdynamics, Inc. v. Firma Wolf, 653 F.2d 93, 96–97 (3d Cir. 1981); Am. Equip. Corp. v. Wikomi Mfg. Co., 630 F.2d 544, 546 (7th Cir. 1980) (explaining that although a consent decree may in rare cases be denied res judicata effect for overriding public policy reasons, such decrees are generally held to bar a party from subsequently relitigating the issues settled by the decree); Schlegel Mfg. Co. v. USM Corp., 525 F.2d 775, 781 (6th Cir. 1975).

56. Dreyfuss, *supra* note 30, at 720; *see also infra* note 115 and accompanying text.

litigation, it may be that the Court no longer values litigation avoidance as highly as it does patent invalidation. Accordingly, there may be little to be gained from settling without the imprimatur of a consent decree.

The patent holder will also likely suffer during the time the case is being litigated. Post-*MedImmune*, there has been considerable debate among practitioners on whether a licensee can recover royalties paid during the time that a case is pending. Some practitioners have suggested that royalties be paid into an escrow account, from which they can be recovered should the patent prove invalid or not infringed.[57] If this is possible, then patent holders will not even be able to use profits from the patent to fund the litigation. Furthermore, as O'Connor points out,[58] if the license is exclusive, then the patent holder will not have the right to negotiate with others. If the license is not exclusive, then in theory, new licensing partners could be found. In practice, however, it is not likely anyone will take—and abide by—a license to a patent that is undergoing challenge.

Of course, the patent holder can charge the licensee more to offset these risks. But that would raise the cost of the products produced under the license and reduce the potential pool of licensees, possibly leading to underutilization of the patented invention.[59] The patent holder could also try to avoid licensee challenges by demanding full payment upfront. However as explained further below, there is a significant downside to that approach. Especially in situations where substantial development of the invention is required before it can be marketed, potential licensees may be unwilling to devote their available funds to paying off licensing fees. Licensees who are just starting out may not even have the money to make the payment. Furthermore, at the time of the negotiation, before the risky business of development and commercialization has been completed suc-

---

57. It has also been suggested that in the future, licensees should bargain for the right to have royalty fees held in escrow during a patent challenge. *See, e.g.*, Mark Henry, *Patent Licensing After* MedImmune v. Genentech, INTELL. PROP. TODAY, Apr. 2007, at 22, http://www.iptoday.com/pdf/2007/4/Henry-Apr2007.pdf; Michael J. Cavaretta & Howard G. Zaharoff, *Patent Licensing Strategies After* MedImmune v. Genentech, IP NEWS (Morse, Barnes-Brown & Pendleton, P.C., Waltham, Mass.), May 2008, http://www.mbbp.com/resources/iptech/newsletters/pdfs/ip0508.pdf; Cooley Alert!, MedImmune v. Genentech*: A Dilemma Removed for Patent Licensees*, COOLEY ALERT! (Cooley Godward Kronish L.L.P., Palo Alto, Cal.), Feb. 15, 2007, http://www.cooley.com/files/tbl_s24News%5CPDFUpload152%5C2554%5CALERT_MedImmuneVGenentech.pdf; Frank X. Curci et al., *Intellectual Property Client Alert: Adapting Your Licensing Strategy to Recent Patent Law Changes*, CLIENT ALERT (Jennings Strouss, and Salmon, P.L.C., Phoenix, Ariz.), June 2008, http://www.jsslaw.com/uploads/publication/Patent%20Licensing%20%206-08%20Client%20Alert.pdf.

58. O'Connor, *supra* note 16, at 447.

59. *See generally* Schlicher, *supra* note 20, at 389–93.

cessfully, the parties may be unable to accurately determine the benefits that will flow from the invention. In contrast, a running royalty provision insures that the licensee will not have to pay if the invention turns out to be worthless. At the same time, it gives the patent holder the assurance of a fair return on a successful invention.

O'Connor has suggested that all of these problems could be solved by compensating the patent holder with a combination of stock and stock options designed to mirror the contribution that the invention makes to the licensee's business. But as he notes, not all patent holders are in a position to administer stock portfolios and comply with the requirements of securities laws; the relationship between the stock price and the performance of the invention may be extremely imperfect (especially if the licensee's business is diversified); and the stock distribution could have an adverse impact on the licensee's other investors.[60] At best, then, this approach can only provide a partial solution to the *MedImmune* problem.

C. Ramifications for Patent Law and Lore

In some ways, it is difficult to understand why the Court left the licensing situation in so much disarray: why Justice Scalia implied, by noting that a promise to pay is not the same as a promise not to challenge, that if there had been a promise not to challenge in the licensing agreement, it would have been enforced;[61] why he thought that there was a genuine contractual dispute concerning the temporal scope of MedImmune's obligation to pay royalties;[62] or why he suggested that there were various aspects of the doctrine of licensee estoppel that *MedImmune* was leaving open.[63] After all, each of these issues appears to have been well settled by three 1960's-era cases. *Brulotte v. Thys Co.*[64] and *Zenith Radio Corp. v. Hazeltine Research, Inc.*,[65] which deemed it misuse to stretch payments beyond the time when the patent was in force or to require payment on account of products that do not use the teachings of the patent, are understood as preventing the patent holder from extracting running royalties for anything other than utilization of a valid patent. And the aforementioned *Lear*, which abrogated the common law doctrine of licensee estoppel, is

60. *See* O'Connor, *supra* note 16, at 451–72.
61. MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 123, 135 (2007).
62. *Id.* at 123–24.
63. *Id.* at 124, 128.
64. Brulotte v. Thys Co., 379 U.S. 29 (1964).
65. Zenith Radio Corp. v. Hazeltine Research, Inc., 395 U.S. 100, 135 (1969).

read as foreclosing the use of no-contest clauses or, indeed, other disincentives to licensee challenges.[66]

In fact, *Lear* appeared to directly address the patent law question at issue in *MedImmune*. It treated the problem of a repudiating licensee as a *harder* question than that of the nonrepudiating licensee. That is, the *Lear* Court started from the assumption that licensees should *not* be permitted to withhold royalties before the patent was invalidated, but came to the conclusion that it would tolerate breach because the prospect of saving money would induce more licensees to mount challenges.[67]

It is, of course, conceivable that the *MedImmune* Court simply failed to understand the lore associated with these cases.[68] But because the Court raised the issue and then made no effort to explain the policy underlying its decision to unmuzzle nonrepudiating licensees, a strong argument can be made that it meant to suggest that licensees need not remain unmuzzled.[69] That is, perhaps the Justices focused narrowly on the procedural issue of what constituted a case or controversy (which, in fact, was never addressed in *Lear*) because they wanted to leave the lore (if not the holdings) of these earlier cases open for reconsideration.

Lear did not, after all, deal with contractual provisions that bar challenges. Although it used broad language that allowed courts and practitioners to believe that provisions barring licensees from challenging patents would be unenforceable, the decision dealt only with the clash between a background rule—the contract-law doctrine of licensee estoppel—and federal patent policy.[70] In a sense, the decision was another manifesta-

---

66. *Lear*, 395 U.S. 653, 673 (1969); *see also* Pope Mfg. Co. v. Gormully, 144 U.S. 224 (1892) (refusing to enforce a licensee's promise not to contest the validity of the licensed patent).

67. *Lear*, 395 U.S. at 673. Indeed, at one point during oral arguments in *MedImmune*, Justice Scalia interjected: "Gee, there's less here than meets the eye." Transcript of Oral Argument at 22, MedImmune v. Genentech, 549 U.S. 118 (2007) (No. 05-608).

68. Certainly, the oral argument was replete with questions that the patent bar has long regarded as decided. *See* Transcript of Oral Argument, *supra* note 67, at 4, 5, 17, 20, 21, 28.

69. *See, e.g.*, MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 124 (2007) ("We express no opinion on whether a *nonrepudiating licensee* is similarly relieved of its contract obligation during a successful challenge to a patent's validity . . . .") (emphasis in original).

70. *Pope*, 144 U.S. at 244. *Pope*, which is also cited for the proposition that no-contest clauses are impermissible, also admits of a narrower interpretation. First, the question presented was whether a court of equity could enforce the contract. *Id.* at 236. Second, the contract did much more than bar the licensee from challenging the licensed patent: among other things, it gave the patent holder the right to control the retail price, it barred the licensee from selling products that competed with the licensed patented inven-

tion of a trope running through a series of contemporaneous Supreme Court cases on the appropriate relationship between federal patent law and state contract law.[71] *Brulotte* and *Zenith* emanated from a similar concern: in both cases, the Court feared that patent holders would utilize contracts to leverage rights under the patent into broader control over the market.[72]

In the new millennium, much of this thinking has changed. The '60s Court had a highly stylized view of the interface between federal and state law: its notion was that anything that was not patented or copyrighted was necessarily in the public domain.[73] But that idea has given way to a much more nuanced appreciation for the benefits of private ordering and for the important role that contracts play in transferring technology and in encouraging creative production outside the formal intellectual property system.[74] Although the Court still occasionally finds a state law that is preempted by federal innovation law and policy,[75] over the forty-odd years since the early cases were decided, the federal/state interface has been substantially refigured. Subsequent to *Lear*, *Brulotte*, and *Zenith*, the Court upheld a contract, negotiated during the pendency of a patent application, which required the payment of royalties even if a patent never issued,[76] it permitted California to protect sound recordings before the Cop-

---

tion (bicycles), and prohibited the licensee from challenging other patents held by the licensor. The Court found the entire license unenforceable as against public policy because it was a restraint on trade. *Id.* at 233–34.

71. *See, e.g.*, *Lear*, 395 U.S. at 668; Compco Corp. v. Day-Brite Lighting, Inc., 376 U.S. 234 (1964); Sears, Roebuck & Co. v. Stiffel Co., 376 U.S. 225 (1964).

72. *See, e.g.*, Zenith Radio Corp. v. Hazeltine Research, Inc., 395 U.S. 100, 135 (1969); Brulotte v. Thys Co., 379 U.S. 29, 33 (1964).

73. *See supra* note 71.

74. There is an extensive literature on the use of contracts to override elements of federal intellectual property law, particularly on the copyright side where the Creative Commons License and the General Public License enjoy considerable popularity. *See, e.g.*, Daniel Laster, *The Secret is Out: Patent Law Preempts Mass Market License Terms Barring Reverse Engineering for Interoperability Purposes*, 58 BAYLOR L. REV. 621 (2006); Lydia Pallas Loren, *Building a Reliable Semicommons of Creative Works: Enforcement of Creative Commons Licenses and Limited Abandonment of Copyright*, 14 GEO. MASON L. REV. 271 (2007); Katherine M. Nolan-Stevaux, *Open Source Biology: A Means to Address the Access and Research Gaps?*, 23 SANTA CLARA COMPUTER & HIGH TECH. L.J. 271 (2007); Molly Shaffer Van Houweling, *The New Servitudes*, 96 GEO. L.J. 885 (2008).

75. *See* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 364 (1991); Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 152–54 (1989).

76. Aronson v. Quick Point Pencil Co., 440 U.S. 257, 266–67 (1979) (Blackmun, J., concurring) (questioning the vitality of *Brulotte* and concurring on the theory that it is important to permit "parties to structure their bargains efficiently").

yright Act covered them,[77] and it approved the application of state trade secrecy law even in cases where the advance could have been patented.[78] Because of these decisions, licensing parties now appear to enjoy considerably more freedom to structure their relationships in the manner that makes the most business sense, so long as their arrangement gives both sides notice of the restrictions imposed and so long as those restrictions yield social benefits.[79]

Equally important, because patents are no longer viewed as monopolies in the true economic sense, the Court has largely abandoned its concerns about leverage. Thus, antitrust and misuse cases from the '60s that found certain practices to be per se violations of the law have largely been overruled in favor of an approach that requires proof that the patent confers market power.[80] In this new environment, the result in *Brulotte* (and perhaps even *Zenith*) is highly questionable. Without knowing how much economic power the patent conferred, there is no way to determine whether the contractual promise to continue payments beyond the term of the patent was produced by leverage. In fact, if the Court were now confronted with a license like the one in *Brulotte*, it would likely start from the assumption that the licensee was paying no more for the invention than it was worth, that compensation was set to reflect the risk that that the patent might be invalidated before it expired, and that the total payout was simply spread over a longer period of time.[81] If the Court needed an economic purpose for the arrangement, two are readily at hand: spreading the

---

77. Goldstein v. California, 412 U.S. 546, 575 (1973).

78. Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 474 (1974). *See generally* Paul Goldstein, *The Competitive Mandate: From* Sears *to* Lear, 59 CALIF. L. REV. 873, 903 (1971).

79. For example, in *Quanta Computer, Inc. v. LG Electronics, Inc.*, 128 S. Ct. 2109 (2008), the court extended the scope of patent exhaustion to process patents and refused to permit parties to a patent agreement to abrogate the patent exhaustion doctrine. In that case, however, the import of the agreement was unclear, producing strong negative externalities. Notably, the *Quanta* Court refrained from explicitly overruling decisions like *Mallinckrodt, Inc. v. Medipart, Inc.*, 976 F.2d 700 (Fed. Cir 1992), where the contract was clear and the bar on reuse protected the quality (sterility) of the inhalers that were the subject of the patent. *But see* Static Control Components, Inc. v. Lexmark, Int'l, Inc., 2008 WL 1970241 (E.D. Ky., Mar. 31, 2009) (No. 5:02-571) (refusing to enforce a contract restriction on exhaustion grounds).

80. *See, e.g.*, Ill. Tool Works Inc. v. Indep. Ink, Inc., 547 U.S. 28 (2006); Charles F. Rule, *The Administration's Views: Antitrust Analysis After The Nine No-No's*, 55 ANTITRUST L.J. 365 (1986).

81. William F. Baxter, *Legal Restrictions on Exploitation of the Patent Monopoly: An Economic Analysis*, 76 YALE L.J. 267, 327–29 (1966); Dreyfuss, *supra* note 30, at 707–12.

payments across a broader time frame reduces the cost of marketing the invention and makes the invention more accessible. Furthermore, the arrangement is in the nature of a loan. It permits a cash-poor licensee to work the patent (for example, to use its teachings to develop new medicines) at lower cost than if the payout term were limited to the life of the patent.[82]

Once it is recognized that patents do not necessarily command significant economic power, and that contractual agreements can serve the public, agreements that discourage (or bar) licensees from challenging the licensed patent look much less suspicious. They avoid the power asymmetry problems noted above and permit licensing by thinly capitalized entrepreneurs and in emerging sectors, such as biotechnology, where the benefits of patented inventions are difficult to evaluate.[83] So long as potential licensees understand at the outset that they will not be allowed to challenge validity after they enter into the contract (that is, so long as the clause is in the contract—as Justice Scalia suggested[84]—rather than arising through a common law estoppel doctrine), they have every incentive to assess patent strength before committing to a license. If a potential licensee were to decide the patent was weak, it could negotiate for a lower payment term or it could simply infringe and wait to be sued. Alternatively, if it wanted to clear its position before investing, the procedural holding in *MedImmune*—if interpreted liberally—would allow it to raise the validity issue in a declaratory judgment action. Indeed, if the pending legislative proposal to institute an opposition proceeding is enacted, the licensee might be able to challenge the patent on the cheap, in the PTO rather than in court.[85]

This approach produces significant social benefits. Indeed, altering the timing of the licensee's propensity to sue arguably produces a better result than the one that obtains under *Lear*. The public acquires free use of the patented information sooner that it would if the licensee were permitted to wait and examine the patent at its leisure. Development and exploitation would then occur in a competitive environment, where they may result in

---

82. For these reasons, courts have begun to question the continuing vitality of *Brulotte*. *See, e.g.*, Zila, Inc. v. Tinnell, 502 F.3d 1014 (9th Cir. 2007) (questioning, but applying, *Brulotte*); Scheiber v. Dolby Labs., Inc., 293 F.3d 1014 (7th Cir. 2002) (questioning, but applying, *Brulotte*).

83. *See, e.g.*, Brief of the Biotechnology Indus. Org. as Amicus Curiae in Support of Neither Party, Quanta Computer, Inc. v. LG Elecs., Inc., 128 S. Ct. 2109 (2008) (No. 06-937).

84. MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 135 (2007).

85. *See supra* note 34.

more—or less—expensive products. In addition, the ability to enter an agreement that locks in royalties furthers the goals of the Bayh Dole Act.[86] Another post-*Lear* development, Bayh-Dole gives universities the right to patent inventions arising out of federally-funded research in order to promote the transfer of fundamental research from academia into the private sector, where it can be used to solve real-world problems.[87] Universities are, however, wary of diverting endowment and tuition funds from research and teaching. Because they instead rely on a steady income from licensing to identify promising academic projects, find appropriate private sector partners, and negotiate agreements,[88] a rule that permits them to contract out of the uncertainties produced by *MedImmune* and *Lear* is particularly advantageous.[89]

Agreements that limit the licensee's freedom to engage in patent litigation also prevent *licensees* from gaming the system. Licensees cannot use the shelter of the patent to develop products, markets, trademark recognition, and a customer base, and then avoid the obligation to pay for the privileges they enjoyed. In fact, as one of us noted in an earlier article, while the *Lear* Court may have been right to be wary of the PTO, its confidence in licensees may have been largely misplaced:

> As long as the patent is thought to be valid, the licensees and the patentee are the only parties with the legal right to practice the invention. The patent confers exclusivity; a successful challenge to it invites competition. Certainly, it is possible that an infringer will enter into competition with the licensee and, unhampered by royalties, undersell the licensee. In that event, the licensee may be inclined to assert its *Lear* right to challenge validity in order

86. 35 U.S.C. §§ 200–12 (2006). The Bayh Dole Act encourages commercialization of university-generated inventions by giving universities the right to retain patents in federally funded research. For the importance of these license in that context, see, e.g., Posting of David Schwartz to the Tech Transfer Blog, Stanford Adds New Clauses to License Agreements in Wake of MedImmune Case, (June 4, 2008), http://www.technologytransfertactics.com/content/2008/06/04/stanford-adds-new-clauses-to-license-agreements-in-wake-of-medimmune-case/.

87. For a discussion of the Bayh Dole Act, see Rebecca S. Eisenberg, *Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research*, 82 VA. L. REV. 1663 (1996); Arti K. Rai & Rebecca S. Eisenberg, *Bayh-Dole Reform and the Progress of Biomedicine*, 66-SPG LAW & CONTEMP. PROBS. 289 (2003).

88. Lita Nelson, From Bench to Bedside: The Changing Landscape of University Technology Transfer in Biotechnology, http://www.ibfconferences.com/ibf/control/presentation/Lita%20Nelson.pdf.

89. *See generally* Paul J. Heald, *Transaction Costs and Patent Reform*, 23 SANTA CLARA COMPUTER & HIGH TECH. L.J. 447, 457–58 (2007).

> to avoid paying royalties. But at this point, the public has little need for an unmuzzled licensee. Judicial review of the patent will occur when the infringer asserts invalidity as a defense to an infringement action brought by the patentee or by licensees who have the right to sue for infringement. Thus, the issue of the patent's validity will be raised no sooner than (and no later than) it would be raised under *Lear*, but without the complications involved in permitting a party to avoid the royalty provisions of a contract entered into after arms'-length negotiation.[90]

To put this another way, when the licensee's freedom to challenge the patent is curtailed, the public exchanges the benefit of an unmuzzled licensee for a different benefit—encouraging licensees to examine the strength of the patents they license at the time they negotiate their agreements. This examination should result in a fairer valuation, and—presumably—lower the price that the public pays for the invention (or its output) from an earlier point in time.

Of course, contractual freedoms are not unlimited. Licensing can certainly be used anticompetitively; the licensee and licensor can enter into agreements that are antithetical to the interests of the public.[91] Accordingly, licenses, particularly agreements that waive public-regarding elements of federal law, must be evaluated with care. Hatch-Waxman settlements are a case in point. In these cases, generic drug manufacturers are paid not to challenge the patents on pharmaceutical products. Lower courts have approved these settlements on the ground that they produce certainty and avoid litigation costs.[92] However, they raise the price of drugs and, hence, contribute to the growing cost of healthcare.[93] Now that *MedImmune* has apparently decided that the interest in challenging patents outweighs the problems associated with proliferating patent litigation, perhaps the case for permitting these settlements should be reevaluated.[94]

---

90. Dreyfuss, *supra* note 30, at 703–05 (footnotes omitted).

91. Examples include the types of licenses that were once considered "No-Nos" by the Justice Department. *See* Rule, *supra* note 80.

92. *See, e.g.*, *In re* Tamoxifen Citrate Antitrust Litig., 466 F.3d 187, 203 (2d Cir. 2006); Schering-Plough Corp. v. F.T.C., 402 F.3d 1056, 1075 (11th Cir. 2005); Valley Drug Co. v. Geneva Pharm., Inc., 344 F.3d 1294, 1304 (11th Cir. 2003) (rejecting antitrust challenges to settlements).

93. *See, e.g.*, *In re* Cardizem CD Antitrust Litig., 332 F.3d 896, 911 (6th Cir. 2003) (finding a settlement to be a per se violation of the Sherman Act).

94. See C. Scott Hemphill, Paying for Delay: Pharmaceutical Patent Settlement as a Regulatory Design Problem, 81 N.Y.U. L. REV. 1553 (2006); see also Harry First, Controlling the Intellectual Property Grab: Protect Innovation, Not Innovators, 38 RUTGERS

Another problem more pertinent to this discussion is informational. Potential licensees may, because of their own business expertise and experience, be in a better position to know about the technology at issue than an examiner in the PTO, but they are not necessarily in possession of full information. Some kinds of information are uniquely in the hands of the patent holder. For example, only the patent holder is likely to have evidence of the true date of invention or whether the patentee was the true inventor.[95] Some third-party evidence, such as evidence of attempts to sell the invention more than a year before the patent application was filed, may be equally difficult to locate.[96] Thus, any examination of the patent at the time the license is negotiated will be incomplete and will therefore not perfectly protect the public interest in a robust public domain. The bottom line is that while *MedImmune* may have cast doubts on the popular understanding of *Lear*, how far it allows parties to order their own relationship remains to be determined.

II. FIVE APPROACHES TO LICENSING IN LIGHT OF *MEDIMMUNE*

To make the question of the continued vitality of the 1960s case lore concrete—and to examine the approaches patentees can take to resolving the problems that *MedImmune* poses—this Part discusses five types of licensing agreements. These arrangements vary in both their distributional and social ramifications. Each attempts to allocate back to the licensee some of the risk of patent invalidation, but as a consequence, each also redistributes the commercial risks associated with the profitability of the invention. Socially, these approaches vary in terms of the degree to which they leave the licensee free (and motivated) to challenge the patent. Approaches that give the licensee less freedom to challenge the patent during the lifetime of the license are in some ways more protective of the public interest than is *Lear* because—unlike the estoppel doctrine—they alert potential licensees to the restriction and thus encourage them to examine the patent prior to negotiating their licenses. The result is that these arrangements foster early challenge and reduce the period in which the public is forced to pay a supracompetitive price for the invention. At the same time,

---

L.J. 365, 390–98 (2007). Indeed, as of this writing, the Senate is considering a bill that would bar these settlements. See supra note 35.

95. *See* 35 U.S.C. § 102(g), (f) (2006). *See, e.g.*, Benitec Austl., Ltd. v. Nucleonics, Inc., 495 F.3d 1340, 1349 (Fed. Cir. 2007). Significantly, in that case the Federal Circuit acknowledged the possibility that the patent was invalid, but held that the case was nonjusticiable even under *MedImmune*.

96. *See* 35 U.S.C. § 102(b) (2006).

however, the more freedom licensees retain to challenge the patent, the more efficacious the challenge, for the licensee gains time to discover information (such as prior art or prior sales) that could be used to invalidate the patent.

How well any of these arrangements survives *Lear* depends, of course, on whether *Lear* and its companion, *Brulotte*, remain good law and on how broadly their mandates are interpreted. These five approaches, which could certainly be modified or amplified by the thoughts of creative licensing executives or business negotiators and their legal counsel, are (a) a paid-up license; (b) a license coupled with a consent decree; (c) an arbitration clause; (d) a clause adjusting royalties to either allow for a challenge or to reflect that a challenge was unsuccessful; and (e) a termination on challenge clause.

Apart from the paid-up license approach, all of these licensing proposals should be understood against the background (evident in the license at issue in *MedImmune*) that licensing agreements usually give the licensee an unfettered right to terminate the portion of the license dealing with patents—as opposed to any provisions dealing with trade secrets or know-how—by simply giving notice. Usually, the licensee also has the right to stop paying royalties for activities under any patent held invalid by a final decision of a court (and not subject to appeal). In some cases these rights are explicitly bargained for, but in many cases they have simply become part of the standard "boilerplate" in typical license agreements. Both provisions provide the licensee some measure of relief if the licensed patent is invalidated by a third party. Accordingly, in most cases, licensees are not, strictly speaking, fully "muzzled."

A. A Paid-up License

Perhaps the most straightforward way for a patent holder to avoid the risks imposed by *MedImmune*, and to regain control over income derived from licensing, is to require the licensee to pay the full licensing fee upfront in a single lump sum payment. If economic conditions permit such an arrangement, the patent holder would secure compensation for the disposition of its rights, while the risk of invalidation and any subsequent decline in the value of the license would be borne by the licensee.

Straightforward as this approach might be, its utility is questionable from both a business and a social policy perspective. Small companies, and especially startups, will often not be in a position to pay out the full cost of the license at the beginning of the relationship. As important, even if the licensee is willing to assume the risk of invalidation, it may not be ready to pay a lump sum and assume all of the business risks associated

with commercializing the invention. For its part, the patent holder may be concerned that the commercial value will be higher than projected, and that it will wind up conferring a windfall on the licensee.

The commercialization problems may, however, be less severe than they appear. As suggested earlier, in some circumstances, a package of stock and stock options can be designed to track the licensee's success utilizing the patented invention. Alternatively, unexpected differences in licensed activity from what the parties calculated at the time the compensation was originally set could be handled by providing for a lump sum payment and then adding two adjustment clauses. One would be an enhancement in payment, which would reflect any increase in the economic value projected at the time the compensation was originally set; the other would provide for a refund to the licensee if net revenues fall below the initially projected amount. These adjustment clauses could also be calibrated to various time periods. For instance, they could provide for additional royalty fee payments or refunds upon comparing the projected amount to the actual amount for some time period less than the life of the license, such as every five years.

As to cash-flow, licensees that lacked the funds to pay the upfront costs might instead provide for payment by negotiable instrument: that is, the licensee would issue a series of negotiable instruments at the time the license was entered into with maturity dates apportioned over the life of the license. If the patent holder needed funds immediately, it could negotiate these instruments without recourse (albeit, most likely at a discount). Furthermore, if the licensee were a small or thinly capitalized company, it could provide notes secured by its assets, including its rights under the license, to make the notes more readily negotiable. Of course, this would require an adjustment in the typical licensing terms to allow the license to be assigned by the holder of the note if there is a default. And it would only make sense if arrangements were made so that if there were a default, there would be only one party holding a security interest in the license able to immediately seize the license or asset. But that could also be arranged: all of the notes could be negotiated together or a single note could include several maturity dates. In either case, the license should also include an acceleration clause for default on any note or installment.

Another alternative is a license providing for a series of payments over time without regard to licensed activities. This would be a firm contractual obligation entered into at the time the license was granted which should be independent of any later determination with regard to the validity of the licensed patents.

These approaches would ameliorate some of the discomfort that both parties might feel in projecting the revenue likely to accrue over the life of the licensed patents, but they too are not without their problems. Depending on the size of the lump sum payment, the patent holder will still run considerable risk that compensation for the invention will fall short of the benefit it confers on the licensee. The licensee will not challenge the patent if actual earnings are lower than projected earnings because it will want its refunds. On the other hand, if the actual earnings are greater than the projected return, the licensee could very well decide to challenge. If the patent is invalidated, it is possible that a court will view the payment adjustment as a species of running royalty, and not require the licensee to continue to pay. Indeed, negotiable instruments, if still in the hands of the patent holder, might also be vulnerable to challenge in that circumstance.

Even a one-shot, lump sum payment will likely be lower than the net present value of the projected running royalty. Typically, the net present value of the projected royalty stream would be calculated using an anticipated discount based on the projected time value of money over the life of the license. The patent holder would have to reduce his price to reflect the benefit to him of this particular arrangement. This "bargain" price would arise from shifting some of the risk of patent invalidity from the patent holder to the licensee: the price would reflect the risk, assumed by the paid-up licensee, that after paying the lump sum, the licensed patent would be invalidated by a third party. The discount would likely be considerable: not only would the licensee have paid for an unnecessary license under an invalid patent but it would be at an economic disadvantage compared to its competitors in the licensed field who had not paid such a lump sum and therefore do not need to recover it in their sales price.

From the perspective of federal patent policy, this arrangement is also somewhat dubious. Because running royalties are the standard way of allocating benefits, paid-up licenses would be utilized only if *Lear* remains good law. But if it is good law, then a paid-up license could be viewed as frustrating the "private attorney general" policy underlying *Lear*. After all, once the licensee has paid for the right to practice the invention, it is less likely to challenge the patent. Indeed, if it is an exclusive licensee, the probability that it will destroy its own competitive advantage is close to zero.

But this problem may also be somewhat more attenuated than it initially appears. Requiring the licensee to pay the entire value of the license upfront provides it with formidable incentives to investigate the patent very carefully. If the potential licensee harbors any doubts about the patent's validity, it will either ignore the patent (which would benefit the public by

lowering its costs of doing business), or it will require a substantial discount in the licensing fee (which would, again, lower the cost of doing business), or—now that *MedImmune* appears to have opened the courthouse door—it will bring a declaratory judgment action (and provide the public with the benefit of free use of the invention). Not only would early scrutiny lead to sooner invalidation of the patent, a party facing the prospect of paying high fees has more incentive to challenge the patent than does a licensee (especially an exclusive licensee).

The paid-up license approach may also be viewed as incompatible with *Brulotte*. That is, a court may be concerned that the economic disadvantages endured by the licensee (loss of startup funds and continuing installment payments) violate the principle that inventions that are not patented are free to use. But as suggested in Part II, there are strong reasons to believe that *Brulotte* is no longer good law. Unless the patent has considerable market power, there is no reason to think that the licensee agreed to pay more than the invention was worth or that it failed to bargain for a discount to account for the possibility that a third party would invalidate the patent. Furthermore, the arrangement has a substantial benefit: a lower fee is attractive, especially to thinly capitalized entrepreneurs and to startup companies.

Nonetheless, the approaches that use negotiable instruments or installments may appear particularly vulnerable to a *Brulotte* challenge because in those cases, it is evident that payments will continue even if the patent is invalidated by a third party. But even here, the benefits of the arrangement are clear. Essentially, these options, which delay the time when payment is due, provide the licensee with a loan to help finance its commercialization efforts. Spreading out the cost of the license over a longer period of time is thus efficient because it reduces deadweight loss.[97] Significantly, since *Brulotte* was decided, the Court upheld an arrangement requiring payments beyond the patent term. In *Aronson v. Quick Point Pencil Company*, negotiations took place while the patent application was pending. Anticipating the possibility that the patent would be rejected, the

---

97. *See, e.g.*, FTC & Dep't of Justice, Antitrust Enforcement and Intellectual Property Rights: Promoting Innovation and Competition 12 (2007). The report concluded that:

> [c]ollecting royalties beyond a patent's statutory term can be efficient. Although there are limitations on a patent owner's ability to collect royalties beyond a patent's statutory term, that practice may permit licensees to pay lower royalty rates over a longer period of time, which reduces the deadweight loss associated with a patent monopoly and allows the patent holder to recover the full value of the patent, thereby preserving innovation incentives.

Id. (citation to *Brulotte* omitted).

license provided for one royalty rate if the patent issued and another if it did not.[98] After the patent failed to issue, Quick Point sued for a declaration that it was not obliged to continue to pay royalties. It lost. The Court rejected the leverage theory that supported *Brulotte*[99] and suggested that the agreement created added incentives to innovate and transfer technology.[100] Finally, it noted that *Brulotte* was not unfair to the public or to the licensee. The public had free use of the invention after it was disclosed.[101] And Quick Point had notice that it would be bound by the agreement even if no patent issued:

> [T]he parties contracted with full awareness of both the pendency of a patent application and the possibility that a patent might not issue. The clause de-escalating the royalty by half in the event no patent issued within five years makes that crystal clear.[102]

In a concurrence, Justice Blackmun even posited that *Brulotte* had been overruled.[103] And both lower courts and the Federal Trade Commission have suggested that it ought to be reconsidered.[104] Notably, although courts have hesitated to ignore *Brulotte*, there is case law to suggest that installment payments remain due even if the licensed patent is invalidated.[105]

B. A License Coupled With a Consent Decree

Although there are ways in which a paid-up license fulfills the goals of *Lear* better than does *Lear*, it suffers from two serious problems. First, there is the information deficit issue noted at the end of Part II. Because

---

98. Aronson v. Quick Point Pencil Co., 440 U.S. 257 (1979).

99. *Id.* at 264–65.

100. *Id.* at 262–63.

101. *Id.* at 263.

102. *Id.* at 261.

103. *Id.* at 266–67 (Blackmun, J., concurring).

104. *See, e.g.*, Zila, Inc. v. Tinnell, 502 F.3d 1014, 1019–20 (9th Cir. 2007) (questioning, but applying, *Brulotte*); Scheiber v. Dolby Labs., Inc., 293 F.3d 1014, 1017–19 (7th Cir. 2002) (questioning, but applying, *Brulotte*); FTC & DEP'T OF JUSTICE, *supra* note 97, at 117; *see also* Michael Koenig, *Patent Royalties Extending Beyond Expiration: An Illogical Ban From* Brulotte *to* Scheiber, 2003 DUKE L. & TECH. REV. 5.

105. *See* Ransburg Electro-Coating Corp. v. Spiller & Spiller, Inc., 489 F.2d 974, 978 (7th Cir. 1973) (holding that installment payments for past activities were due despite invalidation of the underlying patent); *see also* Zenith Radio Corp. v. Hazeltine Research, Inc., 395 U.S. 100, 138–39 (1969) (commenting favorably on the approach in *Automatic Radio Mfg. Co. v. Hazeltine Research, Inc.*, 339 U.S. 827 (1950), that a lump sum payment which was not tied to actual use of the licensed patent was not objectionable).

licensees lack the capacity to acquire information uniquely in the hands of patentees and may also have difficulty finding information in the possession of third parties, an early evaluation of validity is likely to be imperfect. Indeed, one reason that patents never become incontestable is that invalidating information can arise after examination.[106] In some cases (*MedImmune* is an example), the license is negotiated before the patent issues—that is, before the PTO completes its assessment. Second, a general move to paid-up licenses could have a deleterious impact on technological progress. Without running royalties, it is difficult to adjust the price of the license to reflect experience with the invention. Since there is an ever-present likelihood that it is the patent holder's income that will be diminished, incentives to innovate may be reduced.

These problems would, however, be ameliorated if the parties made a running-royalty license part of the settlement of a legal proceeding in which the validity of the patents to be licensed was or could have been put into issue. Typically this would mean that the parties would have to reach a tentative licensing agreement, and then, either the patent holder would bring a patent infringement suit or the licensee would institute a declaratory judgment action. In either event, the licensing agreement would eventually become part of a consent decree terminating the dispute. If the decree were entered by a court that had proper jurisdiction, any challenge to it—including challenge to the validity of the underlying patent—would be a collateral attack, presumably forbidden by res judicata. Accordingly, the consent decree should foreclose challenge by any party to the proceeding or its privities and perhaps even by any party claiming a benefit under the license.[107] Since courts, in our experience, appear to take little interest in the exact terms of these agreements, the parties might even be permitted to agree to running royalties during the entire time the patented invention is worked by the licensee. If the potential licensee were given a forum in which it could fully explore any validity issues in which it had an interest and in which it could enjoy access to the full panoply of discovery devices, including the right to examine the patent holder's files and subpoena third parties, then concerns about lack of information would abate, as would the problems that animated the *Lear* and *Brulotte* Courts. As noted earlier, courts have upheld such arrangements, even against antitrust challenges in the Hatch-Waxman context.[108]

---

106. *Cf.* 15 U.S.C. § 1065 (2006) (defining incontestability of trademark rights).

107. *See, e.g.*, Am. Equip. Corp. v. Wikomi Mfg. Co., 630 F.2d 544, 548 (7th Cir. 1980).

108. *See supra* note 92 and accompanying text.

This arrangement may, however, have limited applicability. First, there is some danger to the patent holder that the licensee will change its mind and proceed with the lawsuit instead of entering into a settlement. After all, the ability to actually use discovery and proceed before the court is what distinguishes the proceedings from sham litigation and gives legitimacy to the consent decree. Since, under *MedImmune*, the patent holder can no longer use a covenant not to sue as a way to have a losing case dismissed,[109] this approach requires the patent holder to assume the risk that the patent will be invalidated at a very early stage in its life, before the patent holder can earn a return on its investment or inventive efforts. This danger is, however, unlikely to be acute. The cost of litigating a case to judgment is extremely high and creates a significant deterrent to walking away from a deal. Unless there was a dramatic change in the commercial landscape after the parties negotiated the tentative license terms, the license is likely to continue to be economically attractive to the licensee.

More important are problems of timing and expense. Given the crowded dockets in most federal courts, procedural skirmishes could lead to unacceptable delay in consummating the license. Moreover, while *MedImmune* took a liberal attitude toward the availability of declaratory judgment actions in patent cases, it did not abolish all standing and ripeness requirements. Accordingly, before either of the parties could bring an action, it would probably be required to demonstrate, at the very least, that the licensee had already engaged in infringing activity or had made substantial preparation to do so.[110] Although the defendant in the action would certainly avoid asserting standing and ripeness defenses, a court concerned with a lack of adversariness could raise the justiciability question on its own motion.[111] And even if the court did not identify the problem, the decree it issued might not be particularly useful to the patent holder because an overly friendly suit—for example, a suit that was settled before the li-

109. *See, e.g.*, Honeywell Int'l Inc. v. Universal Avionics Sys. Corp., 488 F.3d 982, 995 (Fed. Cir. 2007).

110. *See, e.g.*, Prasco LLC v. Medicis Pharm. Corp., 537 F.3d 1329, 1339–40 (Fed. Cir. 2008) (requiring that the plaintiff to show an affirmative act by the patentee demonstrating an intent to sue); *cf.* Janssen Pharmaceutica, N.V. v. Apotex, Inc., 540 F.3d 1353, 1362–63 (Fed. Cir. 2008) (requiring more than speculative fear of harm to establish that the dispute is "definite and concrete").

111. *Cf.* FED. R. CIV. P. 12(h)(3) (subject matter objections can be raised by the court *sua sponte*).

censee actually used its discovery opportunities—will not likely bind sub-licensees and other parties allegedly in privity with the licensee.[112]

The parties could alternatively try to save time and money by settling the case between themselves and then jointly petitioning the court to dismiss the proceedings without seeking a consent decree. Although some courts have enforced no-contest clauses in such agreements when they encounter them as defenses to patent challenges mounted by licensees,[113] not all courts respect them; judges may be particularly reluctant to recognize such clauses if the parties settled before the licensee engaged in discovery.[114] Furthermore, without the imprimatur of the court, it is difficult to know what constitutes a settlement and what is merely the culmination of negotiations to license the patent. Genentech, for example, tried to characterize the license with MedImmune as a settlement.[115] Furthermore, even if a court were inclined to enforce the terms of a settlement against the licensee, it would be highly unlikely to hold anyone else bound by it.

C. An Arbitration Clause

While the consent decree approach may comply with federal patent policy, it utilizes federal court resources in ways that might be considered objectionable—for a trumped up suit that is not designed to produce the public benefit of releasing the invention into the public domain. Instead, the parties could assume the costs of testing the patent by agreeing to arbitrate disputes concerning validity and infringement. Under this approach, the license would be structured to meet the parties' business needs and could include running royalties, but instead of leaving the licensee free to challenge the patent in court as contemplated by *Lear* and *MedImmune*, the licensee would instead be required to arbitrate.

This approach has benefits and drawbacks as both a private and a social matter. On the business side, it rights the imbalance created by *MedImmune*. So long as the parties agree to keep the record of the arbitration—including any evidence presented and any arguments made to the

---

112. *See, e.g.*, Hansberry v. Lee, 311 U.S. 32, 43–44 (1940) (noting that requirements of due process and full faith must be satisfied for a nonparty to be bound); *cf.* Taylor v. Sturgell, 128 S. Ct. 2161, 2173–74 (2008) (rejecting the theory of virtual representation).

113. *See* Dreyfuss, *supra* note 30, at 720; *see, e.g.*, Flex-Foot, Inc. v. CRP, Inc., 238 F.3d 1362, 1370 (Fed. Cir. 2001) (holding that the licensee is contractually stopped from challenging the validity of the patent "[b]ased on the clear and unambiguous waiver of future challenges to the validity of the . . . patent in the settlement agreement").

114. *See* Dreyfuss, *supra* note 30, at 720; *see, e.g.*, Warrior Lacrosse, Inc. v. Brine, Inc., No. 04-71649, 2006 WL 763190, *26 (E.D. Mich. Mar. 8, 2006). *See generally* Cooley Alert!, *supra* note 57.

115. Transcript of Oral Argument, *supra* note 67, at 29.

tribunal—confidential, then even a successful challenge will not invalidate the patent or even provide information to other licensees, potential licensees, or possible challengers.[116] As a result, the parties would be on a more equal footing: the patent holder would not be put into the position of risking the entire value of the patent in the dispute resolution. This arrangement also gives the parties the opportunity to structure the procedures to meet their needs. For example, if the parties are concerned about resources, they could limit discovery and the overall length of the proceeding. Alternatively, a licensee worried about its capacity to obtain sufficient information could bargain for the use of specific discovery tools or reserve the right to utilize judicial proceedings on issues—such as inequitable conduct and the on-sale bar—which are not easily developed in an arbitration proceeding.

Admittedly, there are some disadvantages to arbitration, but they are not particularly severe. For the patent holder, one disadvantage is that arbitration is generally cheaper than litigation (and, as noted above, can be made cheaper still). Accordingly, there is a risk that the reduced cost will make the licensee too ready to challenge the patent. However, because arbitration awards bind only the parties,[117] the arbitration clause will still protect the patent holder from the danger that third parties will benefit from an invalidity decision. Of course, matters submitted to an arbitrator can wind up in court. As discussed previously, the parties can agree to litigate particular issues and one or the other side may seek to overturn the arbitral decision[118] or enforce it. Judicial decisions are presumptively public, but they can be sealed when confidentiality is warranted.[119] For the licensee, the main problem is that the arbitrator may be more deferential to the PTO than judges, making it less likely that a challenge to the patent will win. But, again, the problem does not appear to be acute.

The compatibility of this approach with federal patent policy is, however, somewhat difficult to reckon. Although arbitration makes it less onerous for the licensee to avoid its obligations under the license, the pub-

---

116. Notice that the parties have settled an arbitration must, however, be made public, under 35 U.S.C. § 294 (2006).

117. 35 U.S.C. § 294(c).

118. *See* 9 U.S.C. § 10 (2006) (giving district courts authority to vacate an award that is procured by fraud, corruption, misconduct or bias, or the arbitrators exceeded their powers). *See generally* Karl P. Kilb, Note, *Arbitration of Patent Disputes: An Important Option in the Age of Information Technology*, 4 FORDHAM. INTELL. PROP. MEDIA & ENT. L.J. 599, 619–21 (1993).

119. *See, e.g.*, Commercial Union Ins. Co. v. Lines, 239 F. Supp. 2d 351, 358 (S.D.N.Y. 2002) (holding that a sufficient showing of a need for confidentiality was not made in that case).

lic need not be informed about the nature of the arbitral award. Accordingly, *Lear*'s goal of putting advances that should not be patented into the public domain will be frustrated. On the other hand, the Patent Act explicitly provides that parties may agree to arbitrate disputes regarding validity and infringement; it also provides that the parties are bound by the arbitral award.[120] The availability of binding arbitration for patent invalidity challenges suggests that while *Lear* might require the licensee to be free to defend its own interests in invalidating wrongfully issued patents, federal policy does not always require the licensee to simultaneously protect the public interest by releasing inventions for free public use.[121] Nor does it restrict adjudication of validity to the federal judiciary.

D. A Clause Adjusting the Royalty Either to Reflect or Permit a Challenge

If arbitration is considered insufficient protection for the licensee or the public interest, another solution is to use licensing agreements that include a provision adjusting the royalty rate to reflect validity challenges. One approach is to provide in the license agreement that the licensee may challenge at will, but also provide for an increase in the royalty rate should the challenge be unsuccessful. A second idea is to establish a three-tier system, with the rate increasing once a challenge is mounted, and providing for an even higher royalty if the challenge is not successful.[122] Both arrangements solve the *MedImmune* problem because they increase the risk to the licensee, thus putting it on a more equal footing with the patent holder.

The first approach—a different royalty for patents that have survived challenge—gives the licensee a full and fair opportunity to challenge the patent whenever it finds information suggesting that the patent is invalid. Although the increase in the rate creates a disincentive and may seem in-

---

120. 35 U.S.C. § 294.

121. There are other ways in which federal policy reduces patent challenges. For example, in *Benitec Austl., Ltd. v. Nucleonics, Inc.*, 495 F.3d 1340, 1348–49 (Fed. Cir. 2007), the declaratory plaintiff possessed evidence of invalidity. Nonetheless, the court held the case nonjusticiable under the new standard imposed by *MedImmune*. In some ways, the *Blonder-Tongue* decision has a similar effect. *See generally* Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313 (1971). It discourages suit because it puts the challenger at a competitive disadvantage with respect to everyone else in the field of the invention: that party must pay the full litigation cost of invalidating the patent while its competitors enjoy the outcome for free. As a result, there is an incentive to hold back, to wait and see whether someone else will do the hard work of putting the invention into the public domain.

122. In fact, Stanford University has already announced that it will use this approach. *See* Schwartz, *supra* note 86.

consistent with *Lear*, patents that have survived a challenge are generally perceived by the business community as more valuable than untested patents. Accordingly, the higher rate should be regarded as reflecting the economics of the relationship, rather than as a penalty for challenging the patent. In some ways, it too is similar to the licensing arrangement approved in *Aronson*, where the Court emphasized that the licensee was paying for the value that it had received.[123] Furthermore, the different rate does not entirely eliminate the incentive to challenge: if the challenge is successful, the licensee will terminate the license and escape the obligation to pay royalties. At the same time, the risk of incurring a more costly obligation increases the licensee's motivation to investigate the patent before negotiating a license. As noted earlier, if the potential licensee determines the patent is invalid, it will ignore the patent or insist on a lower licensing fee. Either way, the public benefits from lower costs during the interval before the patent expires or is successfully challenged.

Much the same can be said of the further option, requiring a higher royalty during the time when the case is being adjudicated. Since the pendency of patent litigation can be quite lengthy,[124] this option erects quite a strong disincentive to sue. On the other hand, it also creates a strong incentive to investigate the patent prior to licensing it. In addition, this option is economically justified. Patent litigation is not only lengthy, it can also be very expensive.[125] Imposing an increased royalty during the period of challenge defrays the additional economic burden defending such a challenge imposes on the patent holder. Especially in emerging sectors where many of the licensees cannot afford high fees at the outset of the relationship, this approach is preferable to setting a high royalty rate on all licensees to cover the potential expense of defending a challenge.

There is also another alternative along these lines: a license can be offered at two different rates, depending on whether the licensee wishes to retain the right to challenge the patent. As to *Lear*, the analysis here is much like the one suggested above: the licensee is muzzled by the contract, but has greater incentives to vet the patent prior to licensing. As a

---

123. Aronson v. Quick Point Pencil Co., 440 U.S. 257, 263 (1979).

124. See, e.g., Jay P. Kesan & Gwendolyn G. Ball, How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes, 84 WASH. U. L.R. 237, 282 (2006).

125. *See, e.g.*, R. Polk Wagner, *The Supreme Court and the Future of Patent Reform*, 55 FED. LAW. 35, 36–37 (2008) (observing that for "litigation with less than $1 million at risk, litigation fees from initiation of the lawsuit through appeal were $500,000; with $1 million to $25 million at risk, the fees rose to just over $2 million; and when more than $25 million is at stake, litigation expenses approached $4 million").

business matter, however, this arrangement will only make sense for the patent holder if the license is exclusive. If the license is nonexclusive, then once one licensee opts for preserving the ability to challenge (and agrees to pay at the higher rate), others may well opt for the lower rate, knowing that they can always terminate their agreements in the event that the licensee with the right to challenge does so successfully.

To counter that possibility, the patent holder could couple the lower rate with an obligation to pay for the full term of the patent regardless of any invalidation—that is, with a refusal to permit the licensee to terminate at will. Such a measure could, however, run into an objection based on *Brulotte*. But as noted earlier, there are many reasons to believe that *Brulotte* is no longer good law. Furthermore, just as in *Aronson*, where a dual-rate contract helped the Court understand why, as a general matter, concerns over leverage are misplaced, the differential rate here provides concrete evidence that risk allocation matters; if a licensee is willing to accept the risk of invalidation, it can enjoy the benefits of the patent at a lower rate. In fact, the greater differential between the two licensing rates—that is, the more the royalty is reduced for those licensees who are willing to forgo a challenge—the more credible the argument that there is a social benefit to the arrangement. Indeed, this approach should be of special interest to entrepreneurs at the technological frontier: to startup patent holders and universities who want to lock in an income stream without sacrificing the opportunity to utilize running royalties to allocate the business risks associated with commercialization, and to thinly capitalized licensees, who need to lower the cost of licensing.

E. A Termination-on-Challenge Clause

In some ways, the best way to deal with *MedImmune* is for the patent holder to bargain for the right to terminate the license should the licensee choose to challenge the validity of the patent. With respect to litigation risks, this would fully restore the parties to the pre-*MedImmune* situation. It has the further benefit of harmonizing U.S. licensing practices with those of other countries. Indeed, in Europe, where there are no standing requirements for bringing actions to nullify a patent, termination-on-challenge clauses are routinely used to protect licensors from legal actions on the part of licensees. Significantly, although EU competition guidelines prevent licensors from requiring licensees to forbear from challenging pa-

tent validity, the European Commission has found termination-on-challenge clauses compatible with competition guidelines.[126]

A termination provision must be carefully drafted to maximize the patent holder's options. It should not provide for automatic termination, but, rather, for an option to terminate. This is because there will be cases in which the patent holder will want the license to remain in force. For instance, if the license continues, the patent holder can characterize the case as a contract dispute, and bring the action in state (rather than federal) court.[127] The contract should also be structured to fully protect the patent holder by providing that the option is triggered if the licensee provides direct or indirect aid to a third party in challenging any of the licensed patents. It should also extend to a challenge to any of the licensed patents whether US or foreign, and to a challenge via any official proceeding, be it an opposition, a nullity suit, a reexamination, an interference, or a declaratory judgment action.

If such a provision is contemplated, care should also be taken not to refuse to separately license any unrelated patents included within the license. Otherwise, the licensor may be accused of unfairly protecting weaker patents from attack by bundling them with stronger patents. However, it is quite reasonable to insist that all the members of a patent family be licensed together, since an attack on the validity of one of them would likely impact the validity of the rest of the family—that is, it can be expected that arguments brought against one family member are likely to be pertinent to other members of the same family.[128]

The compatibility of this provision with federal policy is, clearly, the hardest to determine. The *MedImmune* and *Lear* Courts did not consider the issue because until *MedImmune* was decided in 2007, no such clause

---

126. *See* Commission Regulation 772/04, 2004 O.J. (L 123) 5(1)(c); Commission Notice (EC) 2004/C 101/02, 2004 O.J. (C 101), 112, 113.

127. *Lear* was such a case: the action was initially filed in the California Superior Court. *See* Lear, Inc. v. Adkins, 395 U.S. 653, 660 (1969).

128. From the patent holder's perspective, it would be highly desirable to allow for termination if any member of the family of the patent was challenged, even if the challenged member was not part of the licensing agreement. Otherwise, the licensee might be able to avoid the termination provision by taking rights in one set of countries and then challenging the counterpart in a country where the patent is not covered by the license. Although a successful challenge will theoretically leave the licensed patents intact, as a practical matter, a successful challenge will induce others to infringe and reduce the value of the patent holder's assets. At the same time, the more limited the termination clause, the more easily it can be distinguished from the license found unenforceable in *Pope Manufacturing Co. v. Gormully*, 144 U.S. 224 (1892), which involved a direct promise not to bring a challenge at all. *See supra* note 70.

was considered necessary. Rather, it was assumed that licensees who chose to challenge were required to give up the benefits of the license.[129] As Part II suggested, *MedImmune* can certainly be read as intending to take the safeguards offered in *Lear* one step further, to free all licensees to challenge the validity of the licensed patent. But so much has changed since *Lear*, it makes more sense to read the absence of a discussion of this issue in *MedImmune* as an invitation to rethink the received wisdom of *Lear*. The law at the interface between federal and state law has substantially changed since *Lear* was decided and economists and antitrust courts no longer regard patentees as enjoying the leverage necessary to demand concessions that are not economically justified. Furthermore, the need to rely on licensees as private attorneys general has largely waned: there are now new ways to challenge patent validity inexpensively within the PTO.[130] To some extent, even worries about the operation of the PTO—in particular, the "notorious" differences between the law as applied by the patent office and the law applied by the regional courts[131]—have dissipated through the establishment of the Federal Circuit.[132] Certainly, there are reasons to be concerned about that court's jurisprudence. However, unmuzzling licensees will only bring more cases before that court; it will not necessarily improve its decision-making.

This discussion leaves one other option on the table: the no-contest clause for which Justice Scalia appeared to be looking. Certainly, an argument could be made that no-contest clauses should be as enforceable as the approaches discussed above. So long as they are clear, they give the licensee notice that the time to assess the validity of the patent has arrived. Thus, they promote early entry into the public domain of advances that should never have been patented. At the same time, however, the discussion in Part III demonstrates that it is possible to equalize both the litigation and commercialization risks between the parties—that is, to permit challenges while facilitating payment options that enable continuous evaluation of the benefits of the patent—without completely sacrificing the right to challenge the patent. Furthermore, as a key prerequisite for allowing patent holders to extinguish the licensee's right to challenge after a

---

129. *See, e.g.*, Gen-Probe Inc. v. Vysis, Inc., 359 F.3d 1376, 1381 (Fed. Cir. 2004); Cordis Corp. v. Medtronic, Inc., 780 F.2d 991, 993 (Fed. Cir. 1985).

130. *See supra* note 34 and accompanying text.

131. Graham v. John Deere Co., 383 U.S. 1, 18 (1966).

132. At the time of *Lear*, district court patentability rulings were reviewed by regional appellate courts, while the patent office's decisions were reviewed by the Court of Customs and Patent Appeals. In contrast, the Federal Circuit hears appeals from both the PTO and courts. *See generally* Rochelle Cooper Dreyfuss, *The Federal Circuit: A Case Study in Specialized Courts*, 64 N.Y.U. L. REV. 1 (1989).

license is signed, the licensee must have had the ability to challenge the patent before entering into the license. Yet some of the Federal Circuit cases subsequent to *MedImmune* suggest the court is backsliding by once again, requiring the declaratory plaintiff to show an apprehension of suit.[133] If that is so, then it will be important to give licensees some degree of flexibility. After all, *MedImmune* also suggests that there is no strong federal interest in avoiding litigation. As the Hatch-Waxman cases illustrate, there are circumstances where the public interest may be best served by preventing the patent holder from completely muzzling a party in a particularly strong position to challenge the patent.

III. CONCLUSION

In many ways, *MedImmune* was an excellent decision. In striking the Federal Circuit's apprehension-of-suit requirement, the Court put an end to the practice of using patents of dubious quality to stifle competition. With good reason, however, *MedImmune* was greeted with considerable alarm by the licensing community. By permitting licensees to remain in good standing while attacking the validity of the patents they are working, *MedImmune* creates a difficult imbalance in the relative bargaining positions of patent holders and licensees. This asymmetry is likely problematic in all parts of the economy, but it is particularly troublesome in emerging sectors, like biotechnology, where the criteria for patentability are unsettled and many of the participants are thinly capitalized. For start-up licensors and universities, the bargaining asymmetry, coupled with the inability to assure a steady income stream is especially worrisome. And if licensors shift to a system of paid-up licenses, then there will be potential licensees priced out of the market.

But even here, *MedImmune* could represent a blessing, albeit in disguise. The decision could lead courts to revisit *Lear*, *Brulotte*, and the other 1960s cases expressing distrust with state law that touches on innovation policy. *Lear* and *Brulotte* were likely misguided when they were decided and they are certainly out of step with current economic understanding and business practices. Rules that give licensing parties greater flexibility to structure their arrangements can make licensing more efficient, improve public access to new technologies, and enhance incentives to innovate.

This is not, however, to say that *MedImmune* should be read as giving parties full autonomy to structure their relationships at will. By opening

133. *See supra* note 110.

the courthouse door to licensees in good standing, it suggests that avoiding litigation is not a core social goal. Thus, agreements—such as Hatch-Waxman settlements—that restrain competition in the name of saving litigation costs likewise deserve fresh scrutiny.