JIPEL
NYU Journal of Intellectual Property & Entertainment Law

NOTES

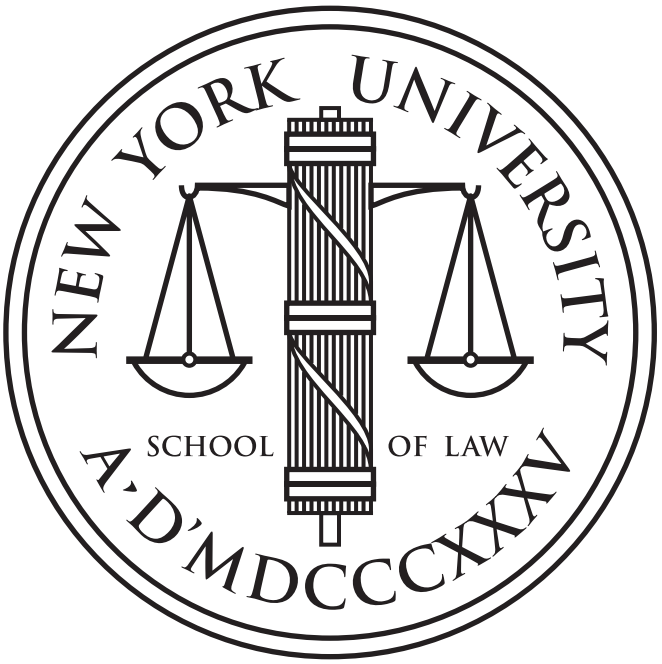
ARTICLES

VOLUME 5 NUMBER 2

NEW YORK UNIVERSITY
JOURNAL OF INTELLECTUAL PROPERTY
& ENTERTAINMENT LAW

VOLUME 5 NUMBER 2

CONTENTS



Statement of Purpose

Consistent with its unique development, The New York University Journal of Intellectual Property & Entertainment Law (JIPEL) is a nonpartisan periodical specializing in the analysis of timely and cutting-edge topics in the world of intellectual property and entertainment law. As NYU's first and only online journal, JIPEL also provides an opportunity for discourse through comments from all of its readers. There are no subscriptions, or subscription fees; in keeping with the open-access and free discourse goals of the students responsible for JIPEL's existence, the content is available for free to anyone interested in intellectual property and entertainment law.

Cite as N.Y.U. J. INTELL. PROP. & ENT. L.

The *New York University Journal of Intellectual Property & Entertainment Law* is published two times per year at the New York University School of Law, 139 MacDougal Street, New York, New York, 10012. In keeping with the Journal's open access and free discourse goals subscriptions are free of charge and can be accessed via www.jipel.law.nyu.edu. Inquiries may be made via telephone (212-998-6101) or electronic mail (law.jipel@gmail.com).

The Journal invites authors to submit pieces for publication consideration. Footnotes and citations should follow the rules set forth in the latest edition of *The Bluebook A Uniform System of Citation*. All pieces submitted become the property of the Journal. We review submissions through ExpressO Bepress (http://law.bepress.com/expresso/) and through electronic mail (submissions.jipel@gmail.com).

A nonpartisan periodical, the Journal is committed to presenting diverse views on intellectual property and entertainment law. Accordingly, the opinions and affiliations of the authors presented herein do not necessarily reflect those of the Journal members.

The Journal is also available on WESTLAW, LEXIS-NEXIS and HeinOnline.

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY & ENTERTAINMENT LAW

PREFACE

Intellectual property law transcends temporal boundaries, forcing the past to keep pace with the developments of the future. The existing legal framework must adapt to new forms of art, data and technology, sometimes overturning existing norms in the name of progress.

The spring issue represents this juxtaposition of old and new through various lenses: Adine Mitrani assesses the protection of rings, historically used for adornment, mystical purposes, as portable bank accounts, and as a signal of socially meaningful codes, in the copyright system. Rachel Scall argues against copyright protection for the burgeoning language of emoji, arguing that such protection will hamper its growth and development. In the patent realm, Julian Pymento criticizes the broadest reasonable interpretation test, arguing that the patent system would benefit from aligning the Patent Trial and Appeal Board claim construction standard with the district court standard. Thomas Cheng sets his sights on another prevailing norm in the patent system, no challenge clauses, exploring the uncharted territory of their antitrust implications. Wen Xue promotes forward-looking obviousness guidelines for the United States Patent and Trademark Office, to encourage uniform results and reduce the level of uncertainty and error in patent prosecution. Finally, NYU Law Professor Christopher Sprigman breaks down the conventions of old in his interview about the Indigo Book, his successful attempt to dismantle the prevailing monopoly of the Legal Bluebook.

On behalf of the JIPEL 2015-2016 editorial board, we sincerely hope you enjoy this issue.

Sincerely,

Caroline Epstein
Editor-in-Chief
NYU Journal of Intellectual Property & Entertainment Law

JOURNAL OF INTELLECTUAL PROPERTY & ENTERTAINMENT LAW AWARD WINNERS 2014-15

BIGGEST CONTRIBUTION
Caroline B. Epstein

BEST NOTE
Florina Yezril

AWARD IN EXCELLENCE
Latore Price
Sherrone Ricks

BEST SENIOR EDITOR
Rebecca Lasky

BEST 2L EDITORS
Hilda Li
Tom Merante

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 5 SPRING 2016 NUMBER 2

OBVIOUSNESS GUIDANCE AT THE PTO

WEN XUE*

* J.D. Class of 2016, New York University School of Law. I am grateful for the guidance from Professor Katherine Strandburg, as well as helpful comments from Professors Rochelle Dreyfuss, Frank Upham, David Kamin, Joanna Langille, members of the Furman Academic Seminar, and members of the Notes Writing Program at NYU Journal of Intellectual Property and Entertainment Law. Any errors are, of course, my own.

INTRODUCTION

The doctrine of obviousness dictates that an invention is patentable only if it is "nonobvious to a person having ordinary skill in the art."[1] This doctrine has the function of balancing the social cost of an exclusive patent right with the contribution that a technological invention brings to society.

The obviousness determination is notoriously indeterminate[2] and is prone to two types of errors. Type II errors, also known as false negatives, refer to granted patents that are actually invalid.[3] These so-called "bad patents" have long been at the center of public discussion.[4] They create deadweight loss and impose various

---

[1] 35 U.S.C. § 103 (2013).

[2] *E.g.*, Gregory Mandel, *The Non-Obvious Problem: How the Indeterminate Nonobviousness Standard Produces Excessive Patent Grants*, 42 U.C. DAVIS L. REV. 57 (2008). For further references, see *infra* notes 24-28.

[3] *See* RONALD E. WALPOLE ET AL., PROBABILITY AND STATISTICS FOR ENGINEERS AND SCIENTISTS 342 (Sally Yagan et al. eds., 8th ed. 2007) ("Nonrejection of the null hypothesis when it is false is called a type II error.").

[4] E.g., FTC, TO PROMOTE INNOVATION: THE PROPER BALANCE OF COMPETITION AND PATENT LAW AND POLICY 5-7 (2003) (arguing that questionable patents deter and raise the cost of innovation, and increase defensive patenting and licensing complications); NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., A PATENT SYSTEM FOR THE 21ST CENTURY 81-82 (Stephen A. Merrill et al. eds. 2004), http://www.nap.edu/catalog/10976/a-patent-system-for-the-21st-century (finding that obviousness standard is too low in granted patents, especially in fields of business method patents and gene sequence patents); John F. Luman III & Christopher L. Dodson, *No Longer a Myth, the Emergence of the Patent Toll*, 18 INTELL. PROP. & TECH. L.J. 12, 13 (2006), https://quote.ucsd.edu/jskrentny/files/2014/08/LumanDodsonPatentTrolls.pdf ("many software patents simply cover inventions that were obvious at the time of the patent application"); Ronald J. Mann, *Do Patents Facilitate Financing in the Software Industry?*, 83 TEX. L. REV. 961, 1026 (raising the criticism that many software patents simply cover inventions that were obvious); David Balto, *Stop Bad Patents Before They Become Problems*, U.S. NEWS, Sept. 5, 2013 (reporting on a campaign to bring attention to patent trolls); Michael D. Frakes & Melissa F. Wasserman, *Does the U.S. Patent and Trademark Office Grant Too Many Bad Patents?: Evidence from a Quasi-Experiment*, 67 STAN. L. REV. 613 (2015) (detailing empirical study confirming that the USPTO is biased in favor of granting patents that are invalid); 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 (2006); John R. Thomas, *Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties*, 2001 U. ILL. L. REV. 305 (2001); Adam B. Jaffe & Josh Lerner, *Innovation and its Discontents: How Our Broken Patent System is Endangering Innovation and Progress*, 6 INNOVATION POL'Y & ECON. 27, 29-31 (2006) (arguing that recent changes in the patent system have resulted in increased litigation and a greater threat of litigation, creating a net social loss); Shubha Ghosh & Jay Kesan, *What Do Patents Purchase? In Search of Optimal Ignorance in the Patent Office*,

costs to society,[5] inflating the market price for products that embody patented inventions, hindering downstream research in areas fraught with substandard patents,[6] and diverting resources to acquire, enforce, maintain, and defend against these substandard patents.[7] Bad patents also contribute to the "patent troll" phenomenon, permitting patent holders to abuse the system by threatening lawsuits with their amassed portfolio of dubious or trivial patents.[8]

On the other hand, Type I errors are patent rejections that should have been granted. Type I errors receive relatively less attention,[9] but are nonetheless important; they reduce incentives in research and development activities, and eventually undermine the patent system's goal of promoting the progress of useful arts.

Errors can occur at both the agency level and the court level. The U.S. patent regime functions in a tiered system. First, in a process called patent prosecution, the applicant files a patent application with the United States Patent and Trademark Office (PTO), which examines the application and makes the decision to grant or reject a patent.[10] Later, a small proportion of granted patents will become the

---

40 HOUS. L. REV. 1219, 1227-35 (2004) (discussing the social costs of low quality patents); T. R. Beard et al., *Quantifying the Cost of Substandard Patents: Some Preliminary Evidence*, 12 YALE J.L. & TECH. 240 (2010).

[5] Beard et al., *supra* note 4 (finding that the economic losses resulting from the grant of substandard patents can reach $21 billion per year by deterring valid research with an additional deadweight loss from litigation and administrative costs of $4.5 billion annually).

[6] *See* Nancy T. Gallini, *The Economics of Patents: Lessons from Recent U.S. Patent Reform*, 17 J. ECON. PERSP. 131, 147 (2002) (noting the negative consequences of lowering the standard for nonobviousness and granting more questionable patents).

[7] *See, e.g.*, Bloomberg West, *TV News Archive*, BLOOMBERG (May 23, 2014, 11:41 PM), https://archive.org/details/BLOOMBERG_20140524_030000_Bloomberg_West?q=4700+3000+patent+troll - start/2516/end/2576 ("Of 4700 patent suits filed in 2012, 3000 were filed by patent trolls"); Balto, *supra* note 4 ("Patents trolls are targeting retailers as easy targets for quick money over questionable patents covering things like Wi-Fi and website features.").

[8] *E.g.*, T.J. Chiang, *What is a Troll Patent and Why are They Bad?*, PATENTLY-O, (Mar. 6, 2009), http://patentlyo.com/patent/2009/03/what-is-a-troll-patent-and-why-are-they-bad.html; Mark A. Lemley & Doug Lichtman, *Rethinking Patent Law's Presumption of Validity*, 60 STAN. L. REV. 45, 48 (2007) ("Sadly, a large and growing number of 'patent trolls' today play this exact strategy, using patents on obvious inventions quite literally to tax legitimate business activity.").

[9] Ron. D. Katznelson, *Patent Reforms Must Focus on the U.S. Patent Office*, MED. INNOVATION & BUS. 77, 78 (2010) ("Allowance errors receive more attention because they are more visible …. Costs of rejection errors are less visible, but no less real.").

[10] 35 U.S.C. § 141(d) (2012) (stating that an applicant of a rejected application have the opportunity to appeal within the agency to the Patent Trial and Appeal Board, appeal to the Federal Circuit); or 35 U.S.C. § 145 (2012) (stating that an applicant dissatisfied with the

subject of patent infringement disputes, and the potential infringer can opt to challenge the validity of the patent either in a federal district court[11] or at the Patent Trial and Appeal Board of the PTO.[12]

This note aims to explore errors at the level of patent prosecution at the PTO.[13] Patent prosecution produces false positives and false negatives and in overall results in a standard of obviousness that is less stringent than at the courts.[14] Some of these errors may never be corrected, such as the final rejection of an otherwise valid patent, or when an alleged infringer takes a license for a patent that is actually obvious. Even when some errors are eventually corrected in litigation, such as when a court invalidates a patent for obviousness, the patent has already been in force for years, imposing significant social costs.[15] Further, these dubious patent grants create *de facto* rights and benefits, setting expectations for future market activities in acquiring, prosecuting, and maintaining patents.

This note proposes that the PTO should proactively promulgate obviousness guidance under its nonlegislative rulemaking authority to elucidate its policy position on obviousness issues long before the court can weigh in. Doing so would

---

decision of the Patent Trial and Appeal Board in an appeal may file a civil challenge in the U.S. District Court for the Eastern District of Virginia).

[11] 28 U.S.C. § 1338 (2011) ("The district courts shall have original jurisdiction of any civil action arising under any Act of Congress relating to patents…").

[12] *E.g.,* 35 U.S.C. § 311 (2012) ("[A] person who is not the owner of a patent may file with the Office a petition to institute an inter partes review of the patent."); 35 U.S.C. § 321 (2012) ("[A] person who is not the owner of a patent may file with the Office a petition to institute a post-grant review of the patent.").

[13] While cognizant of a lively normative debate of the optimal standard of obviousness at the court level, this note does not address this question. Instead, this note hopes to take the court's legal resolution as a given baseline, and analyze the PTO's prosecution errors and biases in relation to the baseline. This is admittedly a partial approach, because my proposal to fix the prosecution bias will affect substantive patent policy, which the courts will later take into consideration when adjudicating validity of granted patents. In other words, the baseline is not exogenous, but rather endogenous to prosecution results.

[14] *See infra* Part I.B. *But see* Mark A. Lemley & Bhaven Sampat, *Is the Patent Office a Rubber Stamp?*, 58 EMORY L.J. 101, 123-24 (2008) (finding that the USPTO is not a rubber stamp but actually rejects a nontrivial amount of applications and that disparities between industries in how they experience patent prosecution may not reflect conventional wisdom about cross-field differences in examination rigor).

[15] See text and citation associated with *supra* note 4. *Cf.* Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 N.W. U. L. REV. 1495, 1502 (2001) (arguing that it may not be cost effective to strengthen patent examination and reduce patents of poor quality because the value of many patents do not depend on their validity).

increase the determinateness of the obviousness doctrine, produce more uniform results in patent prosecution across the PTO, and reduce the level of uncertainty in patent prosecution. Such guidance will also reduce both types of errors and mitigate the problem of a lowered obviousness standard in patent prosecution.

Emerging technology is a particularly apt field for such obviousness guidance. Obviousness in the emerging field of technology suffers the most from doctrinal indeterminateness and bias because legal guidance is the most lacking. It takes time for judicial opinions to catch up with new technology. The PTO, in contrast, can identify emerging technology at a much earlier time than the courts, and thus is at a good position to consider patent policy for emerging technology.

Part I of this note provides the background on the two types of errors inherent in the PTO's decision-making on the obviousness issue. The errors arise from the indeterminate nature of the obviousness doctrine itself, as well as from the patent prosecution procedures at the PTO. The problem is especially salient in areas of emerging technology, where legal guidance and precedents are most lacking.

Part II describes the PTO's authority, particularly its authority to make nonlegislative rules in the form of guidance documents. Although the PTO lacks authority to make legislative rules on substantive patent law, it can still make nonlegislative rules. The PTO has exercised this nonlegislative rulemaking authority at many occasions and promulgated a large number of guidance documents.[16] These guidance documents have not only restated changes in the law that Congress or the courts had already effectuated, but have also provided the PTO's own interpretation of the law and announced its policy position on substantive issues.[17] Still, these previous uses of guidelines have been somewhat backward looking, announcing substantive policies for new technologies long after their advent and usually not until a societal consensus on the policy issue has started to emerge. Forward-looking substantive guidance on obviousness can help improve the determinateness of the obviousness doctrine and counter the effects of bias in patent prosecution.

Part III describes in detail how the PTO should adopt obviousness guidelines in fields of emerging technology. One particular proposal is for the PTO to announce examples of inventions in an emerging field that are deemed obvious, rather than examples of nonobvious, patentable inventions. Part IV argues that obviousness guidelines are feasible under the current regime of institutional

[16] *See infra* Part II.
[17] *See infra* Part II.

authority and will withstand judicial challenges. Part V analyzes the costs and benefits of the proposed obviousness guideline.

I
OBVIOUSNESS ERRORS IN PATENT PROSECUTION

The PTO carries the formidable duty of examining and granting patents.[18] Its vast examiner corps[19] examines and grants over 300,000 patents each year.[20] These patents are at the core of the intellectual property scheme and are highly valued by companies and research institutions. During the examination process, the examiner (an employee of the PTO) and the patent applicant (or her legal representation) engage in a series of back-and-forth oral and written negotiations over whether or not the application meets the standard of obviousness, among other requirements. It takes twenty-seven months, on average, to reach a final disposition on a patent application.[21]

The PTO has often been criticized for granting obvious patents,[22] but errors are often the result of an inadequate system.[23] The highly indeterminate nature of

---

[18] 35 U.S.C. § 2(b)(8)-(13) (2012) (enumerating the various statutory duties of the PTO in advising the President, other executive branches, and congressional committees on intellectual-property-related policy issues).

[19] USPTO, 2014 PERFORMANCE AND ACCOUNTABILITY REPORT 11 (2015), http://www.uspto.gov/about/stratplan/ar/USPTOFY2014PAR.pdf (stating that the PTO employs 9,302 patent examiners as of the end of fiscal year 2014).

[20] *Id.* at 143 (579,782 utility patent applications were filed and 303,931 of them were issued in 2014 alone).

[21] *Id.* at 2 (the average total patent pendency in 2014 was 27.4 months).

[22] *E.g.*, Jaffe & Lerner, *supra* note 4, 32-35, 75, 119-23, 145-49 (criticizing PTO for granting patents on obvious inventions and identifying the realities of the innovation and patenting process); John H. Barton, *Non-Obviousness*, 43 IDEA 475, 477-78 (2003) (arguing that the nonobviousness standard applied by PTO and courts today is not as strict as that articulated by Supreme Court in *Graham*); Matthew Sag & Kurt Rohde, *Patent Reform and Differential Impact*, 8 MINN. J.L. SCI. & TECH. 1, 2 (2007) (noting that "[a]cademics, business leaders, and government officials have all expressed concern that too many patents are issued for [obvious] inventions" (internal quotations omitted)); Carl Shapiro, *Symposium on Ideas into Action: Implementing Reform of the Patent System: Economic Analysis and Critique*, 19 BERKELEY TECH. L.J. 1017, 1018 (2004) (noting that complaints regarding the PTO "typically allege that the [PTO] issues too many questionable patents" including those that were "obvious at the time the patent application was filed").

[23] DONALD A. NORMAN, THE DESIGN OF EVERYDAY THINGS 180-84 (2013) (describing the framework for rules-based mistakes people make in a poorly designed system).

the obviousness doctrine, as well as the procedural characteristics of patent prosecution at the PTO, both give rise to type I and type II errors.

A. Errors Arising from the Indeterminateness of Obviousness

The obviousness inquiry has been described as standard-like,[24] fact-specific,[25] flexible,[26] hard to apply,[27] and indeterminate.[28] What exactly do these terms mean? Consider two proposed rules on traffic safety: Rule A holds that one shall not drive above fifty miles per hour, and Rule B holds that one shall not drive at an unreasonably dangerous speed. People generally agree on what "fifty miles per hour" means, but reasonable minds can differ on what counts as an "unreasonably dangerous speed." The standard of obviousness is akin to the standard of reasonableness; reasonable minds can differ on whether or not an invention is obvious because there lacks a shared understanding of what obvious means.[29] The indeterminateness of the obviousness evaluation stems from this want of a larger core of shared understanding.[30]

---

[24] Michael Burstein, *Rules for Patents*, 52 WM. & MARY L. REV. 1747, 1774-75 (2011) ("[A] pure standard is optimal [for the obviousness doctrine]"). For more general discussions of rules and standards in patent law, *see* John F. Duffy, *Rules and Standards on the Forefront of Patentability*, 51 WM. & MARY L. REV. 609 (2009); John R. Thomas, *Formalism at the Federal Circuit*, 52 AM. U. L. REV. 771 (2003); Rochelle Dreyfuss, *The Federal Circuit: A Case Study in Specialized Courts*, 64 N.Y.U. L. REV. 1, 8-10 (1989) (discussing the "precision" of legal rules and standards).

[25] Pfizer Inc. v. Teva Pharmaceuticals USA, Inc., 803 F. Supp. 2d 409, 441 (E.D. Va. 2011) ("Overall, the court must keep in mind that obviousness is a fact-specific inquiry …."); Jonathan Darrow, *The Patentability of Enantiomers*, 2007 STAN. TECH. L. REV. 2, 8 (2007) ("[O]bviousness is a fact-specific inquiry.").

[26] KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 415 (2007) ("Throughout this Court's engagement with the question of obviousness, our cases have set forth an expansive and flexible approach ....").

[27] Gene Quinn, *When is an Invention Obvious?*, IPWATCHDOG.COM (Feb. 1, 2014), http://www.ipwatchdog.com/2014/02/01/when-is-an-invention-obvious/id=47709/ ("[T]he application of these factors or considerations [of the obviousness inquiry] is exceptionally difficult.").

[28] Mandel, *supra* note 2, at 57 ("[T]he nonobviousness standard … is indeterminate.").

[29] Mandel, *supra* note 2, at 92-95 (arguing that the obviousness doctrine is more indeterminate than the negligence doctrine in torts, and is at a similar level of indeterminacy with the doctrine of obscenity, as in a "you know it when you see it" standard).

[30] *Id.* at 91; Joseph Singer, *The Player and the Cards: Nihilism and Legal Theory*, 94 YALE L.J. 1, 6-7 (1984) (commenting on the problem of legal reasoning being indeterminate and thus on a social level questioning the possibility of setting up a legal system based on the rule of law).

The black letter law of obviousness, as set out in section 103 of the Patent Act, provides that a patent may not be obtained on an invention

> If 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 the claimed invention pertains.[31]

To decide whether or not an invention is obvious, the decision maker—a judge, a jury, a patent examiner, or an administrative patent judge at the PTAB—makes three antecedent findings of fact: the scope and content of the prior art, the difference between the claimed invention and the prior art, and the level of ordinary skill in the art. Based on these three factors, the decision maker reaches a *prima facie* decision on whether or not the difference between the claimed invention and the prior art is obvious to persons having ordinary skill in the art.[32]

After making this *prima facie* decision, the decision maker may take into account so-called secondary considerations to make the final determination of obviousness. All but one secondary consideration are evidence of circumstances before, around, or after the invention that tend to show that the invention is not obvious.[33] Examples of such secondary considerations include evidence that other people have tried to make the invention and failed,[34] or the product implementing the invention has been a commercial success.[35]

Despite these detailed rules and factual findings to guide and structure the obviousness inquiry, the ultimate question of obviousness is, unfortunately, still largely indeterminate. "Obvious" is a subjective term. People with different knowledge bases and creative backgrounds may find an invention obvious or not obvious, yet still be within the reasonable range of judgment. On the one hand, a patent examiner may wrongly grant obvious patents because she likely has lower than ordinary skill in the art. For example, she may not be aware of knowledge that

---

[31] 35 U.S.C. § 103 (2012).

[32] Graham v. John Deer Co., 383 U.S. 1, 17 (1966).

[33] The exception is simultaneous invention, which tends to show that the invention is obvious.

[34] In re Cyclobenzaprine Hydrochloride Extended-Release Capsule Patent Litigation, 676 F.3d 1063, 1081-83 (Fed. Cir. 2012).

[35] Merck & Co., Inc. v. Teva Pharmaceuticals USA, Inc., 395 F.3d 1364, 1376-77 (Fed. Cir. 2005).

is commonly known among people of ordinary skill in the art.[36] On the other hand, the patent examiner may wrongly reject nonobvious patents due to hindsight bias. The examiner often evaluates obviousness of an invention long after the invention has been made, and may find innovative inventions obvious because the solution has already been revealed.[37]

In light of the difficulties in making the *prima facie* obviousness decision, the courts have developed a series of secondary considerations, also known as objective indicia of nonobviousness, to help with the application of the obviousness doctrine. Secondary considerations are factual circumstances that often serve as evidence that the invention was not obvious. These secondary considerations reduce obviousness to an issue of fact-finding based on circumstantial evidence, which the judiciary is accustomed to doing—the most famous example probably being the standard of reasonable care in determining negligence. However, for three reasons explained by Professor Mandel, the obviousness test is more indeterminate than a reasonableness test in negligence. First, negligence is significantly defined by the Hand formula, which provides decision makers with helpful context that is lacking in the obviousness determination. Second, judicial precedents help inform the standard of negligence, but usually do not exist for obviousness determinations. Third, lay decision makers are much more familiar with the perspective of an ordinary reasonable person than that of a technical expert.[38] Furthermore, the Federal Circuit is currently divided on the question of how much weight to afford secondary considerations relative to the *prima facie* obviousness decision. Some panels of the Federal Circuit have held that secondary considerations only come after a *prima facie* determination of obviousness; other panels have placed more weight on secondary considerations.[39]

---

[36] KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 419-21 (2007) (requiring decision makers to take into consideration implicit knowledge shared among persons having ordinary skill in the art even though it is not explicit laid out in prior art documents and rejecting a formalistic conception of that knowledge).

[37] Courts have repeatedly cautioned against hindsight bias. For empirical demonstration of hindsight bias. *See, e.g.*, Gregory Mandel, *Another Missed Opportunity: The Supreme Court's Failure to Define Non-Obvious or Combat Hindsight Bias in* KSR v. Teleflex, 12 LEWIS & CLARK L. REV. 323, 340-42 (2008); Gregory N. Mandel, *Patently Non-Obvious II: Experimental Study on the Hindsight Bias Issue Before the Supreme Court in* KSR v. Teleflex, 9 YALE J.L. & TECH. 1, 18-20 (2007); Gregory N. Mandel, *Patently Non-Obvious: Empirical Demonstration that the Hindsight Bias Renders Patent Decisions Irrational*, 67 OHIO ST. L.J. 1391, 1411-14 (2006) [Hereinafter *Patently Non-Obvious*].

[38] Mandel, *supra* note 2, at 93–94.

[39] *Compare* Alco Standard Corp. v. Tennessee Valley Authority, 808 F.2d 1490, 1504-09 (Fed. Cir. 1986) (Rich, J., dissenting) ("[O]nly commercial success relied on here or below

Judicial definitions add to the shared understanding of obviousness, but they do not make the obviousness doctrine much less indeterminate. The Supreme Court has rejected a rigid application of obviousness in favor of an "expansive and flexible approach."[40] For a long time the Federal Circuit has experimented with a "teaching, suggestion, motivation" test (TSM), which finds an invention nonobvious over prior art references unless there is explicit prior art that teaches, suggests, or provides motivation to combine prior art references. In *KSR v. Teleflex*, the Federal Circuit applied the TSM test and found an invention nonobvious, and thus patentable, when the challenger could not point to a specific piece of prior art evidence that supplied the specific teaching, suggestion or motivation underlying the work. However, the Supreme Court rejected this TSM test when it was applied as a "rigid and mandatory formula," and held that a person having ordinary skill in the art does not necessarily have to be taught or motivated by "published articles and the explicit content of issued patents," but may be motivated by common sense, market demand, or design trend.[41]

Various other judicially developed subtests of obviousness add to the shared understanding of obviousness, but these subtests still leave considerable space for individual discretion. For example, when analyzing new chemical compounds, the Federal Circuit often starts from a "lead compound" and asks whether or not it is obvious to apply certain chemical processes to modify part of its structure.[42] The lead compound approach adds to the meaning of obviousness in the area of chemical compounds, and improves the consistency of the decision making process. However, the doctrine still remains largely indeterminate because the question remains whether or not a certain modification process is obvious to person having ordinary skill in the art.

Judicial precedents of obviousness may shed light on the substantive standard of obviousness and provide some consistency, but their effects have been limited. In a string of cases including *KSR v. Teleflex*,[43] *Leapfrog Enterprises, Inc.*

---

cannot be attributed to Smith's invention as disclosed in his patent but must have been due primarily to other factors.") *with* Arkie Lures v. Gene Larew Tackle, Inc., 119 F.3d 953 (Fed. Cir. 1997) ("The district court's statement that 'secondary considerations are just that – secondary,' suggests a misconception of the role of these considerations in determination of the ultimate question.").

[40] KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 401 (2007).

[41] *Id.*, at 402.

[42] *E.g.*, Pfizer, Inc. v. Teva Pharmaceuticals USA, Inc., 555 Fed. Appx. 961 (Fed. Cir. 2014); Bristol-Myers Squibb Co. v. Teva Pharmaceuticals USA, Inc., 752 F.3d 967 (Fed. Cir. 2014).

[43] *KSR*, 550 U.S. 401.

v. Fisher-Price, Inc.,[44] and *Muniauction. Inc. v. Thomson Corp.*,[45] the courts evaluated inventions that merely updated pre-existing products—a gas pedal for vehicles, a children's learning toy, and municipal bond auctions—with electronics and computer technology, and found them all obvious. These opinions act as the baseline for thousands of other "mere updating" patents, and, to a large extent, render these patents invalid. However, the effect of judicial precedents is limited for two reasons. As courts have also repeatedly held, an obviousness analysis involves a contextual analysis of all the complex and technical facts and does not turn on a single subtest. Even if an invention merely updates a preexisting product and is found *prima facie* obvious, the court may still decide that the invention is not obvious based on evidence of secondary considerations. The question of obviousness defies a rigid definition. Courts strive to conserve the flexible nature of the question, making it a case-by-case determination, and have repeatedly rejected rigid applications of *per se* rules of obviousness.[46]

Understanding the purpose of the obviousness doctrine adds to a shared understanding of the doctrine, but it does not help decision makers determine obviousness with consistency. The obviousness doctrine reflects a utilitarian goal of awarding patents only when necessary to induce innovation. Hence, an invention should be found nonobvious if it, when viewed prospectively, has a low probability of success.[47] Yet, determining whether or not a technical advance is likely to succeed is not much easier than determining obviousness itself.

The indeterminate nature of the obviousness doctrine makes it hard to apply in a consistent way, resulting in two types of errors: granting obvious patents, or rejecting nonobvious inventions. Although one might think such errors are harmless when they cancel one another out and overall produce a standard of obviousness at the right level, the indeterminateness may exacerbate existing

---

[44] 485 F.3d 1157 (Fed. Cir. 2007).

[45] 532 F.3d 1318 (Fed. Cir. 2008).

[46] *E.g., In re* Brouwer, 77 F.3d 422, 425 (Fed. Cir. 1996) ("The test of obviousness … requires that one compare the claim's 'subject matter as a whole' with the prior art"); *In re* Ochiai, 71 F.3d at 1565, 1572 ("[Obviousness] requires a fact-intensive comparison of the claimed process with the prior art rather than the mechanical application of one or another per se rule"); *In re* Baird, 16 F.3d 380, 382, (Fed. Cir. 1994) (rejecting a *per se* rule).

[47] Mandel, *supra* note 2; Robert Merges, *Uncertainty and the Standard of Patentability*, 7 HIGH TECH. L.J. 1, 2 (1992) (stating that the non-obvious requirement "seeks to reward inventions that, viewed prospectively, have a low probability of success."); Michael J. Meurer & Katherine J. Strandburg, *Patent Carrots and Sticks: A Model of Nonobviousness*, 12 LEWIS & CLARK L. REV. 547 (2008) (stating that obviousness should be judged based on whether research project "would be easy or difficult (likely or unlikely to succeed)").

biases that are inherent in the system. For example, when the prospective of receiving a patent as a reward for a research investment becomes more uncertain, risk-averse investors may decide not to finance research and development. This may disproportionately push small firms and individual inventors away from innovative activities. The indeterminateness of the obviousness doctrine may also lead to too many patent grants. Because the number of trivial advances in technology outbalances revolutionary ones, the obviousness doctrine may give rise to more instances of false positives than false negatives.[48] Further, there are generally greater incentives for a patent owner to appeal a wrongly rejected invention than for third parties to challenge a wrongly granted patent. Therefore, false negatives will be corrected while false positives may remain, resulting in too many obvious patents.[49]

For emerging areas of technology, the obviousness inquiry is even more indeterminate.[50] An emerging area of technology is in its initial development stage, meaning that relatively fewer patents have been filed and examined in this field. The general obviousness standard applies to patent applications in emerging areas, but there is little precedent, if any, showing exactly how the obviousness standard is to be applied in the new area. Examiners have limited experience with the technology and patent policy in the new area, making it all the more challenging to apply the obviousness standard consistently.[51]

Emerging technology may further bias the examiner towards lowering the standard of obviousness. An emerging area of technology may be experiencing a rapid growth, resulting in a quickly evolving ordinary skill for active workers in the field. Much of the knowledge may not be reflected in prior art documents, but

---

[48] Mandel, *supra* note 2, at 89-108 (making the same argument).

[49] Joseph Farrel & Robert P. Merges, *Incentives to Challenge and Defend Patents: Why Litigation Won't Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help*, 19 BERKELEY TECH. L.J. 943 (2004).

[50] FTC, *supra* note 4, at 41 ("The PTO recognized that applying patentability criteria to emerging technologies may be difficult or, at minimum, might differ from their application to more established subject matter, and that more senior examiners could assist with the tough judgment calls that ensue.").

[51] FTC, *Hearing on Competition and Intellectual Property Law and Policy in the Knowledge-Based Economy 101*, (Feb. 6, 2002), https://www.ftc.gov/sites/default/files/documents/public_events/competition-ip-law-policy-knowledge-based-economy-hearings/020206ftc.pdf (testimony of Richard Levin, President of Yale University) ("Almost by definition new areas of technology lack well-developed bodies of prior art in earlier patents and in the published literature. This makes it difficult for patent examiners to determine whether a claim meets the required test of … obviousness.").

exists only as implicit knowledge among active workers. It is less likely that an examiner would be in possession of the implicit knowledge in the field. These factors further contribute to the examiner's lack of knowledge,[52] and likely lend to an even lower standard of obviousness.

B. Errors Arising from a Biased PTO's Prosecution Practice

Patent examination is an *ex parte* negotiation between an examiner and the patent applicant. An examiner either allows the patent, or rejects it and explains why it is not patentable. If the examiner rejects a patent, the applicant can argue against it or amend the claims to seek a patent with a narrower scope. The examiner can again grant or reject the patent. This second rejection is usually called a final rejection, but the applicant has an opportunity to amend the claims again, seek an interview with the examiner to persuade her in person, file a continuation, or appeal.[53] It takes on average twenty-seven months for a patent to be disposed.

Patent prosecution serves as a coarse filter to winnow out applications that are obviously invalid. An examiner spends, on average, only about seventeen hours on each patent.[54] The standard of obviousness is likely to be lowered in patent prosecution for reasons such as lack of information, asymmetric incentives to challenge grants and rejections, asymmetric numbers of obvious and nonobvious applications, budgetary incentives, and examiner count incentives.

Examiners' lack of information about prior art lowers the standard of obviousness. Although the examiner conducts the search from an impressive source of prior art, she is not expected to uncover all the relevant prior art, which could include not only documented literature, but also actual products that were used, on sale, offered for sale, or otherwise known to the public. Thus, the examiner may incorrectly conclude that an invention is not obvious because she

---

[52] Melissa F. Wasserman, *The PTO's Asymmetric Incentives: Pressure to Expand Substantive Patent Law*, 72 OHIO ST. L.J. 379, 415 (2011) ("[T]he Agency lacks either a staff with significant knowledge in this technology or the resources necessary to review patent applications in [an] emerging field.").

[53] Mark A. Lemley & Bhaven Sampat, *Examining Patent Examinations*, 2010 STAN. TECH. L. REV. 2, 3 (2009).

[54] Gene Quinn, *High Value Patents—Where Strength Meets Quality*, IPWATCHDOG.COM (Dec. 11, 2014), http://www.ipwatchdog.com/2014/12/11/high-value-patents-where-strength-meets-quality/id=52569/ ("It is unrealistic to expect an examiner to thoroughly review an average of nearly 50 references per patent in the 16 to 17 hours an examiner can spend per patent while processing the necessary number of patent applications." (quoting Stephen Kunin, former Deputy Commissioner for Patent Examination Policy at the USPTO)).

had knowledge of an incomplete set of prior art. The PTO has implemented initiatives to mitigate its lack of information and improve the quality of its work product.[55] The "peer-to-patent" pilot programs, for example, allow third parties to submit prior art to help examiners make obviousness and novelty determinations.[56] Though the pilot programs have proven successful in providing needed information to examiners,[57] they do not promise to eliminate all the problems with lack of information, and it remains to be seen how the PTO can successfully implement them on a larger scale.[58]

Because a patent owner often has a greater incentive to appeal an erroneous rejection than third parties to challenge an erroneous grant,[59] erroneous rejections are more likely to be corrected than erroneous grants. This further contributes to a lowered standard of obviousness in patent prosecution.

The PTO's limited budget and its objective of achieving prosecution efficiency may also create bias towards granting patents.[60] To begin with, patent rejections can be directly appealed in the courts, and judicial review is costly to the

---

[55] *E.g.*, USPTO, *USPTO Launches Second Peer to Patent Pilot in Collaboration with New York Law School*, http://www.uspto.gov/about-us/news-updates/uspto-launches-second-peer-patent-pilot-collaboration-new-york-law-school; USPTO, *Prioritized Patent Examination Program*, http://www.uspto.gov/patent/initiatives/usptos-prioritized-patent-examination-program (explaining that the program gives prioritization and resources to the examination of certain patents upon payment of fees).

[56] USPTO, *USPTO Launches Second Peer to Patent Pilot in Collaboration with New York Law School*, http://www.uspto.gov/about-us/news-updates/uspto-launches-second-peer-patent-pilot-collaboration-new-york-law-school.

[57] *See* Daniel R. Bestor & Eric Hamp, *Peer to Patent: A Cure for Our Ailing Patent Examination System*, 9 NW. J. TECH. & INTELL. PROP. 16 (2010); Christopher Wong & Joseph Merante, *Peer-to-Patent Year One Potential for Implementation in Various Fields of Art Including Biotechnology*, ABA SCITECH LAW, 26, 28 (2008).

[58] *See* Erika Morphy, *New Web Site May Sooth Patent Process*, TECHNEWSWORLD (Mar. 6, 2007), http://www.technewsworld.com/story/software/56129.html?wlc=1294697010 ("The new system also favors large companies that routinely submit patent applications for approval. These firms can maintain staff to monitor the new system and research prior art to shoot down the applications.").

[59] Farrel & Merges, *supra* note 49, at 948-60.

[60] In the 2014 Performance & Accountability Report, the PTO listed as its first goal to optimize patent quality and timeliness. It reported accomplishing providing timely examination of patent applications and reducing backlog of requests for continued examination (RCEs) and patent pendency. USPTO, *supra* note 19, at 41, 46.

agency,[61] so the PTO may be biased towards granting patents to avoid costly judicial review.[62] The pressure on the PTO to reduce its huge backlog may incentivize examiners to reach a final resolution on a patent application. Because the prosecution procedure allows a disappointed applicant to keep her patent alive by keeping filing continuations, the PTO may be incentivized to grant rather than reject borderline patents to reduce the backlog.[63]

Because the PTO is largely self-funded by user fees,[64] it may be incentivized to grant patents so that it can collect issuance fees and maintenance fees. The PTO collects three types of fees from applicants: examination fees that are paid when a patent application is filed, an issuance fee that is paid when a patent is granted, and a maintenance fee that is paid annually permitting the patent to remain enforceable. The examination fees do not cover the actual cost incurred for the PTO to conduct a patent examination. The average cost to examine a patent is about $4,000.[65] The examination fees are only $1,600 for large corporations, and even less for small entities and micro entities.[66] Therefore, the PTO relies on issuance fees and maintenance fees, which are granted only in the event that a patent is allowed.[67] The budgetary concerns may incentivize the PTO, as a whole, to favor patent applicants and over grant patents.[68] This hypothesis is also corroborated in an empirical study, which found that there is a higher allowance rate for patents on

[61] For a third party challenging the validity of a granted patent in courts, the PTO is almost never a party. Wasserman, *supra* note 52, at 406.

[62] *Id.* at 400-417.

[63] *Id.* at 415 ("The backlog of patent applications is a pressing issue to the Agency; the PTO may hope that taking a restrictive stance on patentability standards will result in the filing of fewer patent applications."); *see also* Adam B. Jaffe & Josh Lerner, *supra* note 4, at 136 (2004) (arguing that the current court system incentivizes examiners to "go easy" on applicants and allow their patents).

[64] The PTO has been funded by user fees since 1990. *See* Omnibus Reconciliation Act of 1990, Pub. L. No. 101-508, § 10101, 1990 U.S.C.C.A.N. (104 Stat. 1388), 1388-91.

[65] USPTO, *supra* note 19, at 53 (reporting that total cost per patent production unit was $3,940 in FY 2014).

[66] 37 C.F.R. § 1.16 (2015) (for a utility patent, the basic filing fee is $280, the patent search fee is $600, and the patent examination fee is $720).

[67] 37 C.F.R. § 1.18 (2015); 37 C.F.R. § 1.20 (2015). Wasserman made this exact argument in 2011, *supra* note 52, at 407-09. Although since then fees that the PTO charges patent applicants and patentees have gone up, her conclusion still holds true under today's fee schedule.

[68] Wasserman, *supra* note 52, at 407-15 (arguing that the PTO may favor patent applicants at the expense of the general public under budgetary incentives); Frakes & Wasserman, *supra* note 4 (providing empirical support for the same argument).

technologies with high renewal rates and those filed by large entities, which are the types of patents that earn the most revenue for the PTO.[69]

The PTO's various internal administrative mechanisms translate agency-level biases and preferences into individual applications. These mechanisms, including guidance documents, PTAB decisions, and the Patent Examiner Count System, create systematic bias towards a less stringent standard of obviousness.

Individual examiners' idiosyncrasies may further complicate the result. Specifically, a recent study hypothesized two behavioral patterns that emerge among examiners and lead to bifurcated results in patent prosecution.[70] The Patent Examiner Count System awards examiners "counts," which are used to appraise examiner performance. The Count System awards counts for office actions based on the merits. It also awards counts for disposals of a patent, which include allowance and abandonment of a patent. As a result, an experienced examiner may be incentivized to work quickly to dispose every patent application, getting two points for each application—one for the first office action, and one for the allowance. A junior examiner, who is less experienced and whose work is under more scrutiny, may also take a completely opposite response. She may be incentivized to hold a high patentability standard and continually reject a patent. This could turn an application into a continued source of counts if the applicant is pressed into filing requests for continued examinations (RCEs) and continuations, because the examiner gets counts for a first office action based on the merits, responses to every RCE filed, the first office action in every continuation filed, and allowance or abandonment of the applications.[71] Empirical studies of the PTO's allowance data corroborate this hypothesis, showing that senior examiners take less time on average to prosecute patents to allowance, and have a higher allowance rate than junior examiners.[72] The distinct examiner mentalities are less problematic for a more rule-like doctrine such as novelty, which leaves less space for individual decision maker's discretion, but are more salient in the highly indeterminate doctrine of obviousness, which is highly dependent on the subjective judgment of

---

[69] *See* Michael Frakes & Melissa Wasserman, *Does Agency Funding Affect Decision Making?: An Empirical Assessment of the PTO's Granting Patterns*, 66 VAND. L. REV. 67 (2014) (detailing an empirical study that supports the hypothesis that the PTO's fee structure incentivizes the agency to over grant patents).

[70] Shine Tu, *Luck/Unluck of the Draw: An Empirical Study of Examiner Allowance Rates*, 2012 STAN. TECH. L. REV. 10 (2012).

[71] Tu, *supra* note 70, at 24.

[72] *E.g.*, Mark A. Lemley & Bhaven N. Sampat, *Examiner Characteristics and Patent Office Outcomes*, 94 REV. ECON. & STAT. 817, 821-22 (2012). *See also* Tu, *supra* note 70, at 81.

the decision maker, leading to bifurcated prosecution results based on the luck of the applicant.

II
THE PTO'S BOUNDED AUTHORITY

The obviousness question is ambiguous and indeterminate, leaving a large margin of uncertainty to the examiners to exercise their independent judgment with little guidance. The PTO prosecution practice is set up in a way so that borderline patents are more likely to be granted rather than rejected, for reasons including lack of information, asymmetric incentives to challenge grants and rejections, asymmetric numbers of obvious and nonobvious applications, budgetary incentives, and examiner count incentives. These factors contribute to the perceived phenomena of lowered patents quality.

In the height of societal discontent over bad patents and patent trolls, the courts came to the rescue. The Supreme Court granted certiorari in *KSR v. Teleflex*, found the particular invention obvious and invalid, and announced generally applicable rules in an effort to raise the obviousness standard. The Court rejected the rigid application of the TSM test, and articulated that an invention that merely combines two prior art teachings and yields predictable results is obvious.[73]

As much as the new test seemed a blessing to the then-rampant patent quality problems, it came years after the patents were granted, and the damage was already done in the interim.[74] Why did the test not emerge earlier? It provided more certainty to the obviousness doctrine, especially for emerging technologies, where it is particularly indeterminate how the obviousness doctrine applies and little is known about the level of ordinary skills. It also provided more guidance to the examiner and mitigated inherent biases in the prosecution procedure.

The reason why the test did not emerge earlier is that this is the way that the judiciary works. A court adjudicates the case before it, and only makes rules incidentally when the right case has been brought to it. It has limited means of collecting information that is necessary for rulemaking. A court only has the information that comes before it: factual circumstances of a case, legal arguments made by parties to the case, sometimes *amici* briefs and the court's own research. Overall, a court makes law *ex post*, after the facts have arisen and matured, rather than announcing a general policy from the start.

---

[73] *KSR*, 550 U.S. at 416.

[74] The patent involved in KSR v. Teleflex, US Patent No. 6,237,565 was filed in 2000 and granted in 2001, six years before the Supreme Court invalidated it.

On the other hand, the PTO refrains from making rules because its hands are bound. Under administrative law, the PTO has three potential ways of making substantive rules: informal rulemaking (which is a form of legislative rulemaking) under §553 of the Administrative Procedure Act, nonlegislative rulemaking under the exceptions in §553(b)(3)(A), and announcing rules in adjudications. However, the PTO lacks authority to make legislative rules, and its ability to make nonlegislative rules and conduct adjudications is also subject to limitations.

A. The PTO's Power and Authority

The PTO lacks authority to use informal rulemaking to make substantive rules. As an executive agency, the PTO must act within its congressionally delegated authority. The Patent Act authorizes the PTO to "establish regulations . . . to govern the *conduct of proceedings* in the [PTO]."[75] The Federal Circuit interpreted this statutory language to mean a delegated authority to make *procedural* rules about its proceedings, but not *substantive* rules on patent policy.[76] Thus the PTO is stripped of one of the most powerful agency tools, informal rulemaking (a type of legislative rulemaking), to provide its interpretation of substantive patentability standard.

Because the agency does not hold the power in saying what the law is, its many actions will not receive *Chevron* deference, a high level of deference that courts give to executive agencies' interpretation of statutory language.[77] The PTO only has the statutorily granted authority to issue procedural rules governing patent examination at the PTO, but not substantive rules regarding patentability issues such as obviousness.[78]

---

[75] 35 U.S.C. § 2(b)(2) (emphasis added) ("The Office . . . may establish regulations, not inconsistent with law, which . . . shall govern the conduct of proceedings in the Office[,] . . . shall facilitate and expedite the processing of patent applications[,] [and] . . . may govern the recognition and conduct of agents, attorneys, or other persons representing applicants or other parties before the Office.").

[76] Merck & Co. v. Kessler, 80 F.3d 1543, 1550 (Fed. Cir. 1996) (rejecting PTO's claim for *Chevron* deference in its Final Determination, which interprets Hatch-Waxman Act and Uruguay Round Agreement Acts as limiting the length of potential patent term extensions for patents granted prior to June 8, 1995); *see generally*, Sarah Tran, *Administrative Law, Patents, and Distorted Rules*, 80 GEO. WASH. L. REV. 831, 841-53 (2012).

[77] *Chevron, U.S.A., Inc. v. Natural Resources Defense Council, Inc.*, 467 U.S. 837 (1984).

[78] Merck & Co. v. Kessler, 80 F.3d at 1549 (denying *Chevron* deference to a PTO construction of a patent term extension statute on the ground that "Congress has not vested the Commissioner with any general substantive rulemaking power") (interpreting 35 U.S.C.

It is abnormal in the modern world of expanding executive power that the PTO, an enormous agency commanding over 12,000 employees and charged with the crucial task of examining and granting patents,[79] does not have the authority to make substantive rules. This lack of authority is often attributed to the idiosyncratic power balance in the patent area.[80] Congress established the Federal Circuit as a specialized federal appellate court that has exclusive subject matter jurisdiction on all patent cases. This special court has much experience and expertise in patent policy, and thus retains for itself the authority to interpret the Patent Act and decide substantive patent policy.

Though the PTO can still make rules under its nonlegislative rulemaking authority, this power has been exercised cautiously. Most nonlegislative rules closely follow the language of legislative or judicially made laws. The constrained budget, huge backlog, and criticism on patent quality all push the PTO to avoid aggressive policy positions that may lead to costly litigation when challenged, and focus its limited budget and resources on providing efficient and high quality work on patent prosecution. Compared to other modern executive agencies, the PTO is usually seen as a "weak" government agency that passively follows the Federal Circuit's legal determinations and substantive policy directions, rather than actively engaging in patent policy making.[81]

Congress reoriented the power division in 2011 when it passed the America Invents Act, which established various post grant review proceedings before the

---

2(b)(1)(A)). Obviousness is treated as a question of law and reviewed *de novo*. *Comment Note*, *supra* note 26 (citing In re NTP, Inc., 654 F.3d 1279, 99 U.S.P.Q.2d 1481 (Fed. Cir. 2011)). The line between substantive and procedural rules is blurry, though. And procedural initiatives do affect substantive rights. Programs such as peer to patent and software partnership seek to improve patent quality by expanding the prior art basis for patent examination; fast track patent examination program expedites the issuance of a patent, effectively extending the term of monopoly for a fee.

[79] USPTO, *supra* note 19, at 11.

[80] Melissa F. Wasserman, *The Changing Guard of Patent Law: Chevron Deference for the PTO*, 54 WM. & MARY L. REV. 1959, 2001-02 (2013) (arguing that the role of the Federal Circuit constrains the PTO's executive authority in patent policy). *See generally* Sarah Tran, *Patent Powers*, 25 HARV. J.L. & TECH. 595 (2012) (describing the idiosyncratic power balance between the Federal Circuit and the PTO, and arguing that the America Invents Act of 2011 would bring significant disrupt to the existing power balance).

[81] John M. Golden, *The USPTO's Soft Power: Who Needs Chevron Deference?*, 66 SMU L. REV. 541, 541 (2003) ("[I]n terms of recognized power to speak on substantive questions of law, the USPTO can seem an institutional mite.").

Patent Trial and Appeal Board.[82] These AIA proceedings presented a new, and likely more desirable venue for a member of the public to challenge the validity of a patent.[83] They are faster and less expensive than court litigation.[84] The proceedings are also more favorable to the petitioner; the patent does not receive the presumption of validity, so the challenger only needs to prove invalidity by a preponderance of the evidence, rather than the clear and convincing standard required at court. The public has welcomed the advent of AIA proceedings. From September 2012 to March 2015, 1641 *inter partes* review (IPR) decisions have been instituted.[85] The PTAB has also gone through an aggressive recruitment of administrative patent judges (APJs) to sit on the AIA proceedings.[86]

Most significantly, the AIA granted the PTO the authority to make rules "establishing and governing" AIA proceedings,[87] and the PTO opted for formal adjudication. A panel of three APJs conducts the proceedings in an adversarial fashion. The adjudications allow limited discovery, afford an opportunity for presenting oral arguments, and forbid *ex parte* communications. Scholars have suggested that because Congress delegated the authority to conduct formal adjudications to the PTO, these adjudications should receive *Chevron* deference.[88] This would be a significant power shift because the PTO would be able to make

---

[82] For example, *inter partes* review (IPR) was established under 35 U.S.C. § 311-319, and post-grant review proceedings under 35 U.S.C. § 321-29. (2012)

[83] Anyone but the patent owner can file for an IPR. 35 U.S.C. § 312.

[84] Gene Quinn, *How to Protect Your Patent from Post Grant Proceedings*, IP WATCHDOG (Sept. 21, 2014), http://www.ipwatchdog.com/2014/09/21/how-to-protect-your-patent-from-post-grant-proceedings-2/id=51333/ (pointing out that an IPR typically lasts for a year from the time of institution, and a party spends about $300,000 on it).

[85] USPTO, *AIA Progress* (2015), http://www.uspto.gov/sites/default/files/documents/032615_aia_stat_graph.pdf (showing 2994 petitions have been filed).

[86] USPTO, 2013 PERFORMANCE AND ACCOUNTABILITY REPORT 23 (2014), http://www.uspto.gov/about/stratplan/ar/USPTOFY2013PAR.pdf (identifying "continued aggressive hiring of new [administrative patent] judges"); USPTO, 2012 PERFORMANCE AND ACCOUNTABILITY REPORT 3 (2013), http://www.uspto.gov/sites/default/files/about/stratplan/ar/USPTOFY2012PAR.pdf ("the PTAB boasts of dozens of new administrative patent judges hired from the top echelon of the U.S. intellectual property community"). The PTO hired 61 administrative patent judges in 2014 and 46 more in 2015. USPTO, *supra* note 19, at 55; USPTO, 2015 PERFORMANCE AND ACCOUNTABILITY REPORT 70 (2016), http://www.uspto.gov/sites/default/files/documents/USPTOFY15PAR.pdf.

[87] 35 U.S.C. §§ 316-26.

[88] Wasserman, *supra* note 80 (interpreting the AIA as anointing the PTO as the primary interpreter of core patentability standards, and arguing that courts should afford Chevron deference to PTO adjudication in AIA proceedings).

binding substantive patent rules in AIA proceedings. Yet it is questionable that Congress intended such a significant power shift without clearly indicating it.[89]

Regardless of the highly important question of *Chevron* deference, the AIA sends a clear message that patent quality needs improvement, and the PTO should step up to take action. In a recent paper, Professor Gold highlighted the PTO's soft power, and raised a potentially overlooked point that the PTO already has significant influence over substantive patent policy.[90] He pointed to the successful example of the PTO's utility guideline, and argued that the PTO can influence substantive patent law through nonlegislative rulemaking power by adopting guidance documents.

Following this line of inquiry, this note argues that the PTO can and should make nonlegislative rules on obviousness. The remainder of this section provides a detailed review of the PTO's previous use of nonlegislative rulemaking, supporting the argument that it is within the PTO's power to extend this authority to obviousness guidance.

B. The PTO's Previous Practice of Adopting Guidance

Guidance documents are an indispensable part of the PTO's examination practice. They provide instructions to over 9,000 examiners and the general public on the nuts and bolts of patent examination. The contents of guidance documents include distilling the basis of law from complicated and often contradictory case law, detailing steps and providing flow charts for analyzing the patentability of an invention, and offering form paragraphs for drafting an office action. The PTO routinely provides guidance in the forms of the Manual of Patent Examination Procedures, guidelines,[91] examples,[92] training materials,[93] and forums.[94]

---

[89] *See also* Golden, *supra* note 81, at 545 (expressing "skeptic[ism] that the AIA has worked such a sea change through implicit, rather than express, provision").

[90] Wassermann, *supra* note 52, at 383.

[91] *E.g.*, Examination Guidelines for Determining Obviousness Under 35 U.S.C. 103 in View of the Supreme Court Decision in *KSR International Co. v. Teleflex Inc.*, 72 Fed. Reg. 57,526 (Oct. 10, 2007) [hereinafter 2007 Obviousness Guidelines].

[92] *E.g.*, USPTO, NATURE BASED PRODUCTS EXAMPLES (2014) [hereinafter USPTO, NATURE BASED PRODUCTS], http://www.uspto.gov/patents/law/exam/mdc_examples_nature-based_products.pdf; USPTO, EXAMPLES: ABSTRACT IDEAS (2015) [hereinafter USPTO, ABSTRACT IDEAS], http://www.uspto.gov/patents/law/exam/abstract_idea_examples.pdf.

[93] *E.g.*, USPTO, ANALYZING NATURE BASED PRODUCTS (2015) [hereinafter USPTO, ANALYZING NATURE BASED PRODUCTS], http://www.uspto.gov/sites/default/files/documents/101%20JE%20training%20Nature-Based%20Products%20Module.pdf.

The agency promulgates guidance documents under its inherent authority to make nonlegislative rules.[95] In contrast to legislative rulemaking (also known notice-and-comment rulemaking), which exercises congressionally delegated lawmaking power, nonlegislative rulemaking does not exercise delegated law making power, and is not binding on the public.[96]

Nonetheless, nonlegislative rulemaking is a powerful tool for regulatory agencies to convey information to regulated entities,[97] without the procedures of a notice-and-comment rulemaking.[98] Under administrative law doctrines, nonlegislative rulemaking that affects substantive policy must fall under the one of two exceptions: interpretative rules or policy statements.[99] An interpretative rule provides the agency's interpretation regarding ambiguity in preexisting law. In theory, an interpretative rule does not make new law. When a nonlegislative rule goes beyond a fair interpretation of existing law, it is not an interpretative rule.[100]

---

[94] *See* USPTO, *Subject Matter Eligibility Forum Agenda* (Jan. 21, 2015) [hereinafter USPTO, *Subject Matter Eligibility*], http://www.uspto.gov/patents/law/exam/forum_agenda_20150121.pdf.

[95] Golden, *supra* note 81, at 544; Metro. Sch. Dist. of Wayne Twp. v. Davila, 969 F.2d 485, 490 (7th Cir. 1992) ("All agencies charged with enforcing and administering a statute have inherent authority to issue interpretive rules informing the public of the procedures and standards [they] intend[ ] to apply in exercising [their] discretion." (citations and internal quotation marks omitted)); 1 RICHARD J. PIERCE, JR., ADMINISTRATIVE LAW TREATISE § 6.2, at 306 (4th ed. 2002) ("Any agency has the inherent power to issue an interpretative rule, a policy statement, or a procedural rule to implement a statute it administers."); Thomas W. Merrill, *Judicial Deference to Executive Precedent*, 101 YALE L.J. 969, 1004 (1992) (contending that "once Congress has delegated authority to executive actors under law, the executive agencies must determine what the law means, and need not await a further delegation of interpretative authority from Congress to do so."). *But cf*. Jonathan R. Siegel, *The REINS Act and the Struggle to Control Agency Rulemaking*, 16 N.Y.U. J. LEGIS. & PUB. POL'Y 131, 159 (2013) (suggesting that Congress could explicitly abrogate otherwise inherent authority to issue interpretive rules).

[96] Michael Asimow, *Nonlegislative Rulemaking and Regulatory Reform*, 1985 DUKE L.J. 381, 383 (1985).

[97] Stuart Shapiro, *Executive Discretion and the Rule of Law: Agency Oversight as "Whac-a-Mole": The Challenge of Restricting Agency Use of Nonlegislative Rules*, 37 HARVARD J.L. & PUB. POL'Y 523, 526 (2014) ("Many of these [nonlegislative] policymaking approaches can be characterized as the movement of information from agency managers to other parties.").

[98] 5 U.S.C. § 553(b)(3)(A), (d)(2) (2011).

[99] 5 U.S.C. § 553(b)(3)(A); Robert A. Anthony, *Interpretative Rules, Policy Statements, Guidances, Manuals, and the Like – Should Federal Agencies Use Them to Bind the Public?*, 41 DUKE L.J. 1311, 1323 (1992).

[100] *See* Richard J. Pierce, Jr., *Distinguishing Legislative Rules from Interpretative Rules*, 52 ADMIN. L. REV. 547, 568 (2000).

On the other hand, a policy statement tentatively indicates how agency decision makers will exercise a discretionary power. To qualify as a policy statement the agency document must not bind the regulated entity definitively.[101]

The obviousness standard likely constitutes a policy statement, and may also qualify as an interpretive rule. A policy statement allows the PTO to tentatively indicate how examiners will exercise a discretionary power. In other words, it should preserve the discretionary power for the patent examiners to decide the nonobviousness of each patent application based on the entirety of the facts presented. It announces new law that is tentative and not binding.

PTO has traditionally seen its own role as truthfully applying the substantive patent law as determined by Congress and interpreted by the court. It has carefully avoided the outer boundaries of its authority to avoid legal challenges on its actions, which can be costly both financially and politically. The Federal Circuit may announce precedents that further limit the agency's authority, or disagree with the substance of the PTO policy position, decreasing the credibility and authority of the agency.

Hence, the vast majority of the PTO's promulgated guidance documents track changes of substantive law effectuated by Congress, the Supreme Court, or the Federal Circuit. The obviousness guidelines, promulgated in 2007 and 2010, are examples of guidance documents that merely restate the law without adding much to it. The PTO promulgated the 2007 Obviousness Guideline,[102] in response to *KSR v. Teleflex*,[103] restating the holding of *KSR* and instructing patent examiners to apply seven rationales of obviousness. These rationales were taken from previous Federal Circuit and PTAB case law that are consistent with *KSR*.[104] Three years later, the PTO promulgated the 2010 Obviousness Guideline, restating new developments and interpretations of the Federal Circuit's obviousness doctrine from 2007.[105]

However, the PTO has done more than merely restate the law. Substantive policy making is an inherent component of the PTO's day-to-day function.[106] The

---

[101] Asimow, *supra* note 96.

[102] 2007 Obviousness Guidelines, *supra* note 91.

[103] KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398 (2007).

[104] *See* 2007 Obviousness Guidelines, *supra* note 91, at 57,526-27, 29.

[105] *See* Examination Guidelines Update: Developments in the Obviousness Inquiry After *KSR v. Teleflex*, 75 Fed. Reg. 53,643 (Sep. 1, 2010).

[106] *See* Golden, *supra* note 81, at 543-44 (arguing that the task of patent examination requires the PTO to address patentability questions and thus to "provide answers to substantive questions of patent law") (internal citation omitted); Wasserman, *supra* note 52, at 388-89 ("On a daily

examiners have to make decisions regarding each patent application, including applying the standard of obviousness to the invention. The indeterminateness of the obviousness doctrine means that the decision necessarily involves individual judgment of each examiner. This is especially true in emerging areas of technology, where established definitions of obviousness or legal precedents provide little guidance on the substantive standard. The PTO must make the important determination of obviousness for each patent application. Although the validity of granted patents may later be challenged, the PTO takes a first cut at substantive patent policy, and create *de facto* rights and set expectation for the patentability standards.

The PTO has, from time to time, promulgated guidelines providing its own interpretation of law—in a way "making" law because it adds new content to the previous, unclarified law. These interpretations are in response to new statutes, new judicial opinions, and new facts such as the emergence of new technology. The rest of this section describes the PTO's two different types of such nonlegislative rulemaking: (1) interpretation in response to new changes of law; and (2) guidance as to patentability of inventions in a particular field of technology.

1. Interpretation in Response to Legislative or Judicial Change in Law

In response to three Supreme Court cases on patentable subject matter,[107] the PTO promulgated a series of guidance documents in the forms of guidelines,[108] examples,[109] training materials,[110] and forum presentations at major patent law

---

basis, the PTO must make difficult substantive decisions on issues … such as … standards for nonobviousness.").

[107] Alice Corp. Pty. Ltd. v. CLS Bank Int'l, 134 S.Ct. 2347, 2358-9 (2014), Ass'n for Molecular Pathology v. Myriad Genetics, Inc., 133 S.Ct. 2107, 2120 (2013), and Mayo Collaborative Serv. v. Prometheus Labs., Inc.,132 S.Ct. 1289, 1305 (2012).

[108] *E.g.*, Memorandum, USPTO, Interim Guidance for Determining Subject Matter Eligibility for Process Claims in View of *Bilski v. Kappas* (July 27, 2010), http://www.uspto.gov/sites/default/files/patents/law/exam/bilski_guidance_27jul2010.pdf.

[109] *E.g.*, USPTO, Nature-Based Products, http://www.uspto.gov/patents/law/exam/mdc_examples_nature-based_products.pdf (last visited Mar. 27, 2016); USPTO, Examples: Abstract Ideas, http://www.uspto.gov/patents/law/exam/abstract_idea_examples.pdf (last visited Mar. 27, 2016).

[110] *E.g.*, USPTO, Analyzing Nature Based Products, http://www.uspto.gov/sites/default/files/documents/101%20JE%20training%20Nature-Based%20Products%20Module.pdf (last visited Mar. 27, 2016).

firms.[111] Although a large part of the guidance documents have been devoted to restating the Supreme Court's new rulings and adjusting the PTO examination practice to reflect the changes in law, the PTO also provides its own interpretation of the law and supplements it with specific examples.

Nature-based products is a compelling example of the PTO providing interpretation of law. In 2013, the Supreme Court held in *Myriad* that isolated forms of naturally occurring DNA segments are not eligible for patent protection under 35 U.S.C. § 101.[112] In March 2014, the PTO promulgated a guidance (the March Guidance) to "address the impact of … Myriad" on the "long-standing rule against patents on naturally occurring things."[113] The March Guidance provided the PTO's own interpretation of *Myriad*, and established a three-step process for determining subject matter eligibility of a natural product, which ultimately turned on whether or not the claims as a whole recited something "significantly different" than the natural products.[114] The March Guidance also listed factors that weigh towards and against eligibility,[115] as well as concrete examples of various nature-based products that are or are not eligible subject matter.[116] These factors and examples are not from preexisting statutes or case law but are the PTO's own interpretation of laws.

2. Guidance in Response to Changes in Fact

The PTO has promulgated guidance documents not only to restate changes in the law and fill in legal gaps, but also to more proactively provide guidance in patentability standards regarding new development of facts, such as an emerging field of technology. Perhaps the most famous examples are the utility guidelines promulgated in 2009 and 2011.[117] The utility guidelines have probably been the

---

[111] *See, e.g.*, USPTO, *Subject Matter Eligibility Forum*, http://www.uspto.gov/sites/default/files/documents/101%20JE%20training%20Nature-Based%20Products%20Module.pdf (last visited Mar. 27, 2016).

[112] *Myriad*, 133 S.Ct. 2107.

[113] Memorandum, USPTO, Guidance For Determining Subject Matter Eligibility of Claims Reciting or Involving Laws of Nature, Natural Phenomena, & Natural Products 1 (Mar. 4, 2014) [hereinafter USPTO, *March Guidance*] (internal citation omitted), http://www.uspto.gov/patents/law/exam/myriad-mayo_guidance.pdf.

[114] *Id.*

[115] *Id.* at 4-5.

[116] *Id.*

[117] Examples of comments on the successful implementation of the utility guidance to effectuate policy changes include Burstein, *supra* note 24; Arti Rai, *Growing Pains in the Administrative State: The Patent Office's Trouble Quest for Managerial Control*, 157 U. PENN. L. REV. 2051, 2053-54 (2009); Golden, *supra* note 81.

PTO's most aggressive nonlegislative rulemaking to influence substantive patent policy. They were adopted in response to the emerging technology of gene sequencing and were used to reject patent applications on expressed sequence tags (ESTs).

The gene patent controversy started in the early 1990s when Craig Venter and the National Institute of Health (NIH) filed hundreds of patents on human gene sequences.[118] The patentability of gene sequence has been a topic of considerable debate. One specific issue was the patentability of a particular type of gene sequence known as expressed sequence tags (ESTs). An EST is a short nucleotide sequence that represents a fragment of a cDNA clone. It is an important research tool and can be used as a probe to isolate the full cDNA sequences or as a marker to locate a particular gene on a chromosomal map,[119] but EST's practical benefit is limited to research purposes.[120] Professors Eisenberg and Merges, in an influential opinion letter published in 1995, concluded that patents on ESTs as research tools are undesirable because they discourage subsequent research.[121] They concluded that ESTs likely do not meet the utility requirement because they lack practical or specific utility,[122] and are also vulnerable to obviousness challenges.[123]

In 1999, the PTO promulgated a utility guideline,[124] which specifically rejected patentability for ESTs because of their lack of utility. The 1999 Utility Guideline interpreted utility as "at least one ... credible, specific, and substantial [utility]."[125] The guideline also provided a specific example, stating that a claim to a polynucleotide whose use is disclosed simply as a "gene probe" or "chromosome marker" would not be considered specific in the absence of an explicit DNA target

---

[118] *See* Robert Cook-Deegan & Christopher Heaney, *Patents in Genomics and Human Genes,* 11 GENOMICS & HUMAN GENETICS 383 (2010) (describing the early history of patents on gene sequences and ESTs).

[119] Rebecca S. Eisenberg & Robert P. Merges, *Opinion Letter as to the Patentability of Certain Inventions Associated with the Identification of Partial DNA Sequences,* 23 AIPLA Q.J. 1, 18 (1995).

[120] *Id.* at 18.

[121] *Id.* at 52.

[122] *Id*. at 3-19.

[123] *Id.* at 52.

[124] USPTO, REVISED INTERIM UTILITY GUIDELINES TRAINING MATERIALS (1999) [hereinafter USPTO, UTILITY INTERIM GUIDELINES], http://www.uspto.gov/sites/default/files/web/menu/utility.pdf; Press Release, USPTO, USPTO Offers Training Materials for Interim Written Description and Utility Guidelines (Mar. 1, 2000), http://www.uspto.gov/news/pr/2000/00-15.jsp.

[125] USPTO, UTILITY INTERIM GUIDELINES, *supra* note 124, at 5.

disclosure; therefore, ESTs do not possess specific utility and are not patentable.[126] The PTO later reaffirmed this interpretation in its 2011 Utility Examination Guidelines.[127]

These utility guidelines were widely recognized as successful examples of PTO nonlegislative rulemaking, because the Federal Circuit later explicitly endorsed the PTO's position on patentability of ESTs. An inventor, Dane Fisher, filed a patent application on an EST gene sequence, which was rejected in 2004.[128] On appeal, the Federal Circuit not only affirmed the PTO's finding that Fisher's claim lacked specific and substantial utility,[129] but also explicitly acceded to the PTO's utility guidelines. It "[took] judicial notice" of the guidelines, and stated that the PTO's standards "comport with this court's interpretation."[130] The Federal Circuit also noted that the example in the utility guidelines "is applicable to the facts here," and that the MPEP "particularly explains that [an EST claim] directed to be useful as 'gene probe' or 'chromosome market,' as is the case here, fails to satisfy the specific utility requirement."[131]

Compared with other guidelines that merely restate or interpret case law, the utility guidelines are more forward-looking. They were adopted proactively, at a time when the patentability policy for the emerging technology of ESTs was not completely settled.

The utility guidelines, however, also take on a somewhat backward-looking approach. The gene patent controversy first started at the beginning of the 1990s, public awareness on gene sequencing soared in the mid-1990s, and Professors Eisenberg and Merge's opinion letter was published in 1995, all long before the utility guidelines were adopted in 1999 and 2001. The utility guidelines were more likely codifying the emerging public consensus on EST patentability, rather than the PTO's effort to actively lead and shape the policy.

The PTO's cautious use of guidance documents reflects the agency's modest assessment of its institutional role in the patent regime; it does not seem keen on providing direct input to shape substantive patent policy. The PTO appears to

---

[126] *Id.*
[127] Utility Examination Guidelines, 66 Fed. Reg. 1,092 (Jan. 5, 2001).
[128] *See generally ex parte* Fisher, 72 USPQ.2d 1020 (B.P.A.I. 2004).
[129] *In re* Fisher, 421 F.3d 1365, 1369-70 (Fed. Cir. 2005).
[130] *Id.* at 1372.
[131] *Id.* at 1372-73.

perceive its major task as examining patents efficiently,[132] and with high quality.[133] With regard to policymaking, the PTO views itself as a modest part of a bigger executive mechanism. It focuses on providing information to the executive and legislative branches to facilitate policymaking,[134] rather than proactively shaping substantive policy through its examination practice.

Patent examination is not a mechanical task of merely following the law, however. Especially for questions of obviousness, examiners continually confront new facts, new inventions, and must determine the bounds of the law with each patent application. When the PTO deems a patent nonobvious, it necessarily provides its own interpretation of substantive patent law, making substantive policy determinations. The PTO could leave the highly indeterminate obviousness question to each examiner's individual judgment and passively wait for the judiciary to review and make the ultimate legal determination. Alternatively, the PTO can provide internal training to examiners to standardize prosecution practice, improve consistency and reduce errors. Ideally, the PTO would make its guidelines known to the public to further reduce regulatory cost.[135]

The remainder of this note proposes and analyzes a more forward-looking use of PTO guidance on the doctrine of obviousness in areas of emerging technology. This approach would convey the PTO's policy position to examiners and the public at an earlier time, reducing both type I and type II errors in obviousness determinations, and reducing the structural bias towards a lowered obviousness standard in the PTO's examination procedures.

III
A PROPOSAL FOR OBVIOUSNESS GUIDANCE

The PTO should promulgate forward-looking obviousness guidelines. A particularly good example would be in an emerging area of technology, before a societal consensus is formed over the optimal policy on the emerging technology. Thus, the PTO could provide instructions to examiners and the public about how it

---

[132] It measures efficiency with patent pendency, budgetary expense, and backlog reduction. *See* USPTO, *supra* note 19, at 44-56.

[133] It measures patent quality through reviews of written work product by supervisory examiners and the PTO's customers. *See* USPTO, *Patent Quality Assurance*, http://www.uspto.gov/patent/laws-and-regulations/examination-policy/patent-quality-assurance (last visited Mar. 22, 2016).

[134] *See generally* USPTO, *supra* note 19.

[135] *See infra* Part V.

would determine the obviousness of inventions in a new field of technology. Such instructions would reduce uncertainty and inconsistency in the PTO's examination practice. They would also enable the PTO to play a more active role in the U.S. patent policy regime and to use its expertise and unique position to impact patent policy in a more proactive and affirmative way.

The obviousness guideline would meet the following three conditions. First, it should provide more consistency and reduce indeterminateness in the application of the obviousness doctrine. Second, the guidance should be made within the PTO's authority of nonlegislative rulemaking. Third, the benefits of implementing the guidance should outweigh the costs. Though the discussions in Part III take into consideration all three conditions, the second and third conditions will be further elaborated in Parts IV and V, respectively. The PTO is the first government actor to detect a rising trend of innovation and patent applications in a particular technology field, and is thus positioned to identify an emerging field of technology at the outset. The PTO necessarily faces the question of patentability for these patent applications as it receives them. An obviousness rule is most feasible when there are a sufficiently large number of inventions with the same identifiable core of operative facts.[136]

We have seen examples of such an identifiable core of operative facts in the past. For example, when electronics technology was relatively new, many manufacturers applied the then-nascent electronics technology in existing products, such as vehicle acceleration pedals or children's learning toys, in ways that were predictable.[137] Similarly, when gene-sequencing technology was relatively new, researchers applied DNA sequencing technology to known proteins with known amino acid sequences.[138] Though the courts eventually found the inventions in these examples likely to be obvious, this position was not adopted until over a decade later. In doing so, the courts invalidated many patent grants and disrupted expectations. The PTO has the opportunity to consider the obviousness standards at a much earlier stage, voice its opinion, and proactively influence substantive patent law.

---

[136] *See infra* Conclusion.

[137] *See e.g.*, KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398 (2007) (finding a vehicle acceleration pedal that applies electronics technology obvious); Leapfrog Enter. v. Fisher-Price Inc., 485 F.3d 1157 (2008) (finding children's learning toy that applies electronics technology obvious); Miniauction Inc. v. Thompson Corp., 532 F.3d 1318 (2008) (finding an online auction method obvious).

[138] *See e.g., In re* Deuel, 51 F.3d 1552 (Fed. Cir. 1995); *In re* Kubin, 561 F.3d 1351 (Fed. Cir. 2009).

The PTO also has the capacity and expertise to make the right obviousness guideline. First of all, in the above-mentioned historical lessons of electronics technology and gene sequencing technology, the PTO erroneously made the initial decisions to grant the patents, not because it lacked the expertise to understand the technology, but because it was bound by the rigid TSM test. Now that the TSM test is no longer in place,[139] the obviousness doctrine is free from its constraint, and the PTO is can make its obviousness determinations.

The PTO is also an expert agent competent to decide the standard of obviousness.[140] The PTO as an executive agency in general hires over 9,000 examiners with technical expertise and APJs who are experienced patent attorneys with both legal and technical expertise. The Federal Circuit judges have expertise in patent law, and some also have a technical background, but most judges at the district court or the Supreme Court levels are generalist judges and lack technical expertise. In addition, the PTO has the expertise to understand economic implications of patent policy at a broader level. In 2010 the PTO established a Chief Economist Office to develop its expertise in economic analysis,[141] whereas federal judges do not necessarily have an economics expertise. The PTO also has administrative authority to conduct studies and solicit comments from stakeholders regarding the consequences of a proposed guideline, while the court has to focus on the specific case and the parties involved. The PTO examines hundreds of thousands of patents a year, and has access to the full picture of patent grants and rejections, while only thousands of patent litigations are filed with courts each year.[142] The PTO also has more resources and expertise than the courts to make an obviousness guideline.

---

[139] *See generally KSR*, 550 U.S. 398.

[140] *See* Burstein, *supra* note 24 (arguing that the case for the PTO to make rules for substantive patent law is at least as strong as other executive agencies).

[141] USPTO *Office of Chief Economist*, http://www.uspto.gov/about-us/organizational-offices/office-policy-and-international-affairs/office-chief-economist/chief (last visited Mar. 22, 2016) ("The OCE initiates and oversees groundbreaking economic analysis in the field on the topics of intellectual property protection and enforcement, with the object of fulfilling the USPTO's statutory obligation to provide the President ... and the administration with advice on intellectual property policy."); Arti Rai, *supra* note 117, at 2054.

[142] *See* Gene Quinn, *Patent Litigation Statistics: 1980-2010*, IPWATCHDOG.COM (Aug. 2, 2011), http://www.ipwatchdog.com/2011/08/02/patent-litigation-statistics-1980-2010/id=17995/ (showing patent litigation statistics from 1980 to 2010); Gene Quinn, *The Rise of Patent Litigation in America*, IPWATCHDOG.COM, (Apr. 9, 2013), http://www.ipwatchdog.com/2013/04/09/the-rise-of-patent-litigation-in-america-1980-2012/id=38910/ (showing patent litigation statistics from 1980 to 2012).

The PTO should decide the timing of the obviousness guideline. Part of the PTO's goal in adopting an obviousness guideline is to convey its purported examination position to the examiners and the public *ex ante*, proactively influencing the substantive obviousness standard. However, the PTO may also want to wait until a better time, for example, when more patents with similar facts are filed, for public opinion and known consequences on obviousness policy to develop more fully, or for the agency to learn more about each stakeholder's position.

The PTO can conduct studies and solicit public opinions to gather needed information.[143] Such information could include information germane to the obviousness determination, such as the level of ordinary skill in the art and the implicit knowledge of practitioners in the field. It would also be helpful to consider information relating to the underlying economic justification of patents, such as the cost of development, market projections, and evidence regarding private parties' incentives to engage in a particular line of research and development. The earlier the PTO wishes to make the obviousness guideline, the more costly it will be to collect the information necessary to make predictions about the consequences of a proposed guideline.

The content of the obviousness guideline can take the form of specific examples. Such an example should comprise the operative facts of an invention, relevant prior art, and the level of ordinary skill, as well as the conclusion that the invention *is obvious* in light of the prior art. The guideline should not provide examples of inventions that are *not obvious* because such examples would be inefficient. The standard of obviousness necessarily rises with the expansion of human knowledge; what was not obvious at an earlier time becomes obvious later.

To illustrate, consider a hypothetical obviousness guideline in response to the emerging area of electronics technology in the 1990s.[144] Such a guideline could include one or more specific examples of inventions that apply the electronics technology on, for example, a children's toy, which uses electronic components to generate sounds.[145] The example would describe Prior Art A, which is a children's toy that uses an electro-mechanical record player to generate sounds. It would

[143] 35 U.S.C. § 2(b)(11) (2006).

[144] I use past technology as examples for the purpose of illustration because the nature of the obviousness inquiry makes it hard to discuss technology that may emerge in the future in much detail. Here we also have to assume that the TSM test does not exist because it would be inconsistent with the current flexible, fact-specific obviousness doctrine and preclude the PTO to make the obviousness guideline.

[145] *See* Leapfrog Enter. v. Fisher-Price Inc., 485 F.3d 1157 (2008).

describe Prior Art B, which includes the knowledge of modern electronic technology. The guideline would then reason that the invention in question is obvious in light of the prior art A and B. This example would provide future examiners with a baseline for examining similar patent applications that adapt an old product using new technology in a commonly understood way.

The obviousness guideline could further articulate a generalized rule on top of specific examples. Articulation of a generalized rule would add to the shared understanding of obviousness and further improve the certainty and consistency in obviousness determinations.

Continuing with the above hypothetical, the guideline with electronics technology could further articulate the rationale for finding the specific example of children's toy obvious. The rationale could be, for example, that the invention merely adapts an old product using new technology in a way that is "commonly available and understood."[146] Alternatively, the rationale could be that the invention merely adapts an old product using technology in a way that is "predictable."[147] The choice of different rationales would be a policy decision, as it may result in different obviousness standards. "Predictable" may indicate a less stringent standard of obviousness than "commonly available and understood," because a new technology may be "commonly available and understood," but may not produce "predictable" results.

The obviousness guideline, however, should still leave enough discretion for the examiner, as required by the contextual and flexible nature of the obviousness doctrine.[148] The examiner is required to take the invention as whole and to consider its specific factual circumstances. Therefore, even for a future patent application that fits closely with the obviousness guideline, it is not necessarily obvious. The examiner is still free to consider other factors, such as evidence of commercial success of the invention and prior art "teaching away" from combining the old product with the new technology.

The specific examples and the generalized rule both add to the shared understanding of obviousness in the particular field, but they impose different costs in rulemaking. Specific examples are relatively easy to find, as the PTO can take examples from patent applications. A generalized definition is harder to come up with because it must be based on analysis of a large number of specific examples

[146] *Leapfrog*, 485 F.3d at 1168.
[147] Eisai Co. Ltd. v. Dr. Reddy's Labs. Ltd., 533 F.3d 1353, 1359 (2008) (analyzing *KSR*).
[148] *See supra* Part I.

that share the same core of operative facts. Moreover, it has to be articulated in language that suits different specific examples in the same field. Therefore, an obviousness guideline with a generalized rule is more expensive to make, or has to be made later in time, than an obviousness guideline with only specific examples.[149]

The PTO should also consider the formality of the guideline. The language that the PTO adopts in making a guideline affects how the court will later interpret it. For example, an obviousness guideline should articulate the law in language that closely follows prior case law.[150] The PTO can use recent PTAB decisions as specific examples in the guideline, to show that that it reflects no change in law.[151] Like previous PTO guidelines, the obviousness guideline should also proclaim that it is nonbinding, does not create any legal rights or benefit,[152] and that rejections of patents would be based on substantive law, and it is these rejections that are appealable.[153]

IV
JUDICIAL CHALLENGE TO THE OBVIOUSNESS GUIDELINE

When the PTO promulgates an obviousness guideline to cut back on the patentability of a field of new technology development, a disappointed applicant whose application has been rejected is likely to appeal to the Federal Circuit. There are three ways that the applicant may appeal. First, she may appeal the rejection of the patent itself. Second, she may challenge the guideline, claiming that it exceeded the scope of the PTO's authority; that is, it does not fall within the scope of nonlegislative rulemaking. Third, she may challenge the guideline, arguing that the content of the guideline is not a correct interpretation of the obviousness standard. An obviousness guideline would likely withstand all three challenges.

---

[149] A generalized definition also adds more to the shared understanding of the legal term of "obviousness" than a specific example, and thus adds more consistency to the application of the obviousness standard to later patent applications.

[150] *See* Animal Legal Defense Fund v. Quigg, 932 F.2d 920, 927 (Fed. Cir. 1991) (finding that the PTO's Notice on patentable subject matter did not exceed the PTO's authority because it was "merely interpretive of prior decisional rule").

[151] *See id*. at 927-28.

[152] Note that the guideline should claim that it does not create any *private* legal rights or benefits, but it does create significant *public* benefits, as is the main point of this note.

[153] *See e.g.*, 2007 Obviousness Guidelines, *supra* note 91, at 57,526.

A. Challenge of a Patent Rejection

Though a disappointed applicant may challenge the rejection of her patent, this probably will not affect the validity of the obviousness guideline. In the hypothetical articulated in Part III, the proposed guideline would describe specific examples of an obvious invention, and explain the underlying rationale that an invention is likely to be obvious if it "merely adapt an old product using new technology in a way that is commonly available and understood." Because of the flexible nature of the obviousness doctrine, future applications would still be analyzed on a case-by-case basis, rather than by categorical determinations. The examiner would make the decision whether or not the rationale applies to a later invention based on how similar the facts of the invention in question are to the facts of the example in the guideline. Therefore, even if a disappointed applicant challenges the rejection of her patent at the Federal Circuit, and the Federal Circuit reverses, it only means that the guideline would not control in that particular case, not that the guideline itself was invalid.

B. Challenge of the PTO's Rulemaking Authority

A party affected by the guideline could challenge it on the grounds that the guideline exceeds the PTO's delegated authority to promulgate procedural rules under 35 U.S.C. § 2(b)(2)(A).[154] The party could also claim that any legislative rulemaking by an agency requires the notice and comment procedure under 5 U.S.C. § 553(b), (c).[155] These challenges would likely fail because the obviousness guideline would likely qualify as a nonlegislative rule, as either a policy statement or an interpretative rule under 5 U.S.C. § 553(b).[156]

The PTO's nonlegislative rulemaking authority has rarely been challenged in the past.[157] In *Animal Legal Defense Fund v. Quigg*, the Federal Circuit distinguished legislative rules from nonlegislative rules under the test of "whether

---

[154] Merck & Co., Inc. v. Kessler, 80 F.3d 1543 at 1550 (Fed. Cir. 1996) (interpreting 35 U.S.C. § 2(b)(2)(A) as conferring the PTO the authority only to promulgate procedural, not substantive rules).

[155] 35 U.S.C. § 553 (2006).

[156] Courts have used the terms "non-legislative" and "non-substantive" to describe agency rule making that is exempt from notice and comment procedures under § 553(b)(A) of the APA. In this paper I will use only "nonlegislative" to avoid confusion with the substantive-procedural distinction in *Merck.*

[157] The Federal Circuit has only heard challenges like this twice before, in *Animal Legal,* and *Mikkimeleni v. Stoll*. 932 F.2d 920; 410 Fed. App'x 311 (Fed. Cir. 2010). Both times the Federal Circuit rejected the challenges and affirmed PTO's guidelines.

or not the rule effects a change in existing law or policy which affects individual rights and obligations."[158] In *Animal Legal Defense Fund*, the PTO promulgated a Notice that "[the PTO] now considers nonnaturally occurring non-human multicellular living organisms, including animals, to be patentable subject matter." Before the Notice, the Supreme Court held in *Diamond v. Charkrabarty*,[159] and the Board of Patent Appeal and Interference held in *Ex parte Hibberd*[160] and *Ex parte Allen*,[161] respectively, that live, non-naturally occurring living microorganism, multi-cell plants, and polyploid oysters were eligible for patent protection. The Notice synthesized these three pieces of case law, and expanded the scope of non-natural product to all non-naturally occurring living organisms. Although the Notice expanded the literal reach of the law, the court nonetheless found that it was nonlegislative rulemaking because it represented no change in the law and was merely interpretative.[162]

Under *Animal Legal Defense Fund*, courts will likely look at the substantive effect a nonlegislative rule has over the regulated entities and the extent to which the rule affects individual rights and obligations. Consider the utility guidelines discussed in Part II, which were promulgated at a time when the societal consensus on the issue of ESTs patentability started to emerge.[163] If, hypothetically, the same utility guidelines were promulgated at an earlier time when applicants first started to file patent applications on ESTs, they might fail if challenged under *Animal Legal Defense Fund*. In the hypothetical situation, the PTO might want to argue that the rule did not change the patentability standard, but merely clarified the correct standard and reduced the margin of uncertainty. In other words, it did not affect individual rights and obligations because it merely rejected patents that should not have been granted anyway. However, the court would have likely found the rule constitutes too much of a change to the law because it categorically denies patentability to a whole field of inventions, in a way that unsettles parties' expectations. Thus the guidelines would have significant implications on parties' rights and benefits, and thus could not be made as nonlegislative rules. In reality, the PTO promulgated the utility guidelines after patent examination practice and public consensus had adjusted parties' expectations. Therefore, the rule reflected not a radical change of law, but rather a codification of the new norm.

---

[158] *Animal Legal Defense*, 932 F.2d at 927.
[159] Diamond v. Charkrabarty, 447 U.S. 303 (1980).
[160] *Ex parte* Hibberd, 227 U.S.P.Q. 443 (B.P.A.I. Sept. 18, 1985).
[161] *Ex parte* Allen, 2 U.S.P.Q.2d 1425 (B.P.A.I. Apr. 3, 1987).
[162] *Animal Legal Defense Fund*, 932 F.2d at 927.
[163] *See supra* Part II.

The proposed obviousness guideline would likely qualify as nonlegislative rulemaking under the rationale in *Animal Legal Defense Fund*. The aforementioned specific examples of obvious inventions would not change the law, but merely explain how the obviousness doctrine applies to specific cases, and thus would not affect individual rights and obligations. The generalized rule would be articulated in the context of the specific examples, allowing an applicant to draw distinctions between the facts in her patent application and those in the guidelines. She could make nonobviousness arguments based on the full scope of the facts of her invention. The guidelines would not change the fact-specific, flexible nature of an obviousness inquiry, and thus would not have a substantive impact on the rights of each applicant.

Under general principles of administrative law, the obviousness guideline could also fall under the exception of a policy statement because it would tentatively indicate how agency decision makers will exercise a discretionary power, but would not definitively limit the examiners' discretionary power in future examinations.[164] The obviousness doctrine requires the examiner to make the obviousness determination for each individual patent application based on all the facts of the application. If the guideline were to definitively limit examiners' discretion, it should have been rejected on its substance for being a *per se* rule of obviousness.[165] The obviousness guideline should preserve the flexibility of the obviousness inquiry,[166] and examiners should be genuinely open to make obviousness determinations in future examinations.

The obviousness guideline could also be an interpretative rule that construes the statutory language of "obvious" because it does not adversely impact the patent applicant's rights and obligations. The guideline would not change the flexible, fact-specific nature of each obviousness determination. Although some patent applicants would be refused the issuance of a patent, any alleged right to a patent never existed in the first place. The guideline would merely clarify the PTO's interpretation of how the obviousness standard applies to the new technology, and does not reflect a change of substantive obviousness standard. Losing a patent right that the applicant would not have deserved in the first place would not constitute an adverse impact of the patent applicant's right in this situation. The obviousness

---

[164] A policy statement is a nonlegislative rule that tentatively indicates how agency decision makers will exercise a discretionary power. *See e.g.*, Simmons v. ICC, 697 F.2d 326 (7th Cir. 1982); Am. Bus Ass'n v. United States, 627 F.2d 525 (D.C. Cir. 1980).

[165] See citations associated with note 46.

[166] *See generally* Asimow, *supra* note 96, at 393.

guideline would also likely deter some patent applications from being filed at all, but the inventors would still be free to file a patent and appeal a patent rejection to the Federal Circuit.

C. Challenge of the Substantive Content of the Guideline

Because the obviousness guideline qualifies as nonlegislative rulemaking, it is likely not reviewable for lack of finality or ripeness. For an agency action to be final, it must be the "consummation of the agency's decision making process," and must be one "by which rights or obligations have been determined, or from which legal consequences will flow."[167] The obviousness guideline would likely not be final because it provides a tentative opinion as to the weight of a particular fact pattern, without determining any legal rights and obligations or creating any direct legal consequences as to the patentability of any application. The obviousness guideline is similarly unlikely to be ripe because it does not directly address regulated entities' conduct. It would not produce direct consequence on patent applicant's action or create any hardship, but would merely clarify the content of the law as applied to a particular fact pattern.[168]

Even if the Federal Circuit decided to review the substance of the obviousness guideline, the PTO's guideline would likely be persuasive under the four factors that the Supreme Court used in Skidmore v. Swift to analyze the persuasiveness of agency's interpretation: (1) the thoroughness of the agency's investigation; (2) the validity of its reasoning; (3) the consistency of its interpretation over time; and (4) other persuasive powers of the agency.[169] Because the guideline would take the narrow scope of listing specific examples of obvious inventions, under the first two *Skidmore* factors, the court would likely look at the facts and reasoning supporting the examples. The court may also look at policy considerations such as the nature of the innovations, the content and scope of patent applications, the potential downstream effect of granting a patent right, and the likely incentivizing effect of a patent right. The PTO would likely be able to provide sufficient support and reasoning regarding the obviousness of specific invention examples.

The third *Skidmore* factor is the consistency of the PTO's interpretation over time. Under this factor, the court would likely consider how the new guidelines

---

[167] Mikkilineni v. Stoll, 410 F. App'x 311, 313 (Fed. Cir. 2010) (citing Bennett v. Spear, 520 U.S. 154, 177-78 (1997)); Mark Seidenfeld, *Substituting Substantive for Procedural Review of Guidance Documents*, 90 TEX. L. REV. 331, 375-76 (2011).

[168] Seidenfeld, *supra* note 167, at 381.

[169] Skidmore v. Swift & Co., 323 U.S. 134, 140 (1944).

substantively differs from pre-existing law, and how the new guidelines had been promulgated and drafted with sufficient reference to previous case law. The obviousness guideline would likely be considered consistent with pre-existing law because they would merely apply the obviousness standard to specific examples, reflecting an incremental, rather than radical, change.

The court would also likely be influenced by public opinion and secondary sources, which also take time to develop and settle. When a field of new technology is in its early development stage, the public lacks sufficient knowledge to analyze the appropriate policy. For the PTO to prematurely promulgate a rule before a public consensus has started to settle, even if the rule eventually turns out to be acceptable, it runs the risk of judicial reversal. However, the obviousness guideline would not be a broadly applicable hardline rule that definitively affects future applications. Rather, the guideline would be based on specific examples, and reasons why these examples are obvious. The controversy on the patentability of the whole area of technology is unlikely to affect the fact-specific obviousness determination regarding the specific examples in the guideline.

The obviousness guideline would be narrow in scope because it would purportedly only address the obviousness of specific examples in the guidelines, and would not substantively affect future patent applications in the field. Therefore, it is within the PTO's ability to draft well supported, and well-reasoned nonlegislative rules that do not reflect an abrupt change in the law of obviousness. Furthermore, it is likely that the court would agree with the content of the guidelines should it chooses to review the guideline on the substance.

V
COST-BENEFIT ANALYSIS OF THE OBVIOUSNESS GUIDELINE

Part V analyzes the potential initial cost of promulgating the guidelines and the ensuing benefits that the guideline will confer on post-guideline adjudications.

A. The Cost

If an agency chooses to promulgate an obviousness guideline, it has control of the timeline. A typical notice-and-comment rulemaking takes about six to twelve months to promulgate. The obviousness guidelines would probably takes a shorter time because nonlegislative rulemaking does not require all the formalities of legislative rulemaking.

The PTO would want to support its position with economic study results and public feedback. These measures are costly, but they help the PTO make the right decision in deciding the content of the guidelines. Moreover, they buttress the guideline's persuasiveness when under judicial review. Public comments have provided helpful inputs and influenced the PTO's nonlegislative rulemaking in the past. For example, the PTO promulgated a Patentable Subject Matter Guidance in March 2014, and after public comment, promulgated another Patentable Subject Matter Guidance in December 2014.[170] In the December Guidance, the PTO incorporated changes from the March Guidance in response to public comments.[171] For example, the December Guidance evaluates whether or not the claimed product is "markedly different" from a naturally occurring counterpart not only in its "structural difference," as articulated in the March Guidance, but also in its "function, or other properties."[172]

B. The Benefit

The obviousness guideline is economically beneficial because it reduces the two types of errors in patent examination, counters the lowered obviousness standard in the examination practice, and reduces the number of unworthy patent applications and therefore the number of unworthy patents.

First, the obviousness guidelines would reduce the indeterminateness of the obviousness doctrine and increase the consistency of obviousness determinations

---

[170] USPTO, *March Guidance*, *supra* note 113; 2014 Interim Guidance on Subject Matter Eligibility, 74 Fed. Reg. 74,618 (Dec. 16, 2014) [hereinafter December Guidance].

[171] For all the comments regarding the March Guidance, see USPTO, *Public Comments on Guidance For Determining Subject Matter Eligibility Of Claims Reciting Or Involving Laws of Nature, Natural Phenomena, & Natural Products* (Jul. 2, 2014), http://www.uspto.gov/patent/laws-and-regulations/comments-public/public-comments-guidance-determining-subject-matter. The December 2014 Guidelines also eliminated the multi-factor test in the March 2014 Guidelines and adopted a more streamlined process.

[172] 2014 Interim Guidance on Patent Subject Matter Eligibility, 79 Fed. Reg. 74,619, 74,621, 74, 623 (Dec. 16, 2014). For examples of public comment in response to the guidance, see AARP, *Re: Guidance for Determining Subject Matter Eligibility of Claims Reciting or Involving Laws of Nature, Natural Phenomena, & Natural Products*, at 3, http://www.uspto.gov/sites/default/files/patents/law/comments/mm-a-aarp20140731.pdf (last visited Mar. 27, 2016); American Bar Association, *Re: Comments of the American Bar Association Section of Intellectual Property in Response to the USPTO's Guidance for Determining Subject Matter Eligibility of Claims Reciting or Involving Laws of Nature, Natural Phenomena & Natural Product*s, at 14-19, http://www.uspto.gov/sites/default/files/patents/law/comments/mm-a-abaipl20140731.pdf (last visited Mar. 27, 2016).

in patent examination. Greater certainty in the obviousness doctrine would reduce instances of both false positives and false negatives.

This phenomenon is demonstrated in Figure 1 below. Suppose before the obviousness guideline, the margin of uncertainty is defined as (–a, a) on the axis of a level of obviousness. An examiner would always reject a patent if the level of nonobviousness is below –a, and always grant a patent if the level of obviousness is above a. Inventions with a level of obviousness between –a and a are uncertain, and it is possible that the examiner may grant it or reject it. The obviousness guideline would add to the shared understanding of obviousness and narrow the margin of uncertainty to (–b, b). An examiner would now reject a patent if the level of nonobviousness is below –b, reducing instances of false positives that may happen in the range of (–a, –b) absent the obviousness guideline. Similarly, an examiner will now grant a patent if the level of nonobviousness is above b, reducing instances of false negatives that may happen in the range of (b, a).

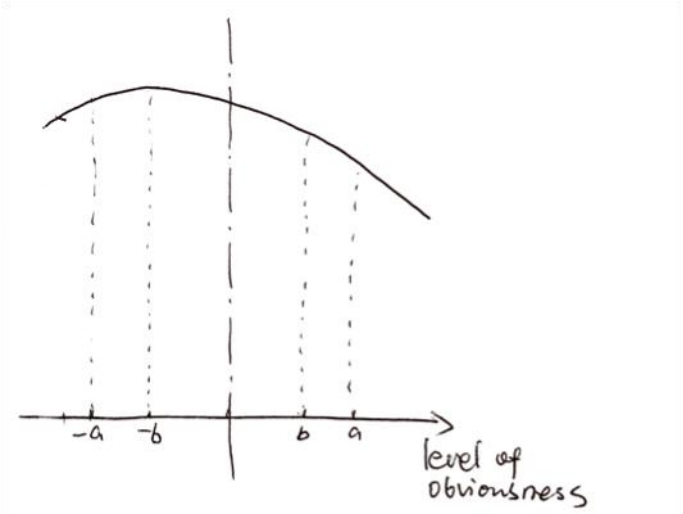


Figure 1: The x-axis indicates level of obviousness; the y-axis indicates distribution of patent applications. The distribution is merely illustrative and is not based on actual data. However, to reflect the fact that there are more minor improvements than major improvements, the curve is skewed towards the left side.

Second, the obviousness guidelines would counter the structural bias in the PTO's examination practice of applying a lowered standard of obviousness, further reducing instances of false positives. Consider Figure 1. It is uncertain whether an inventor should reject or grant a patent with this margin of uncertainty. Suppose

she would grant a patent when on weighing all the evidence, she believes the patent is patentable beyond a level of certainty, say 40%, because she is biased towards granting a patent. We have an effective bar of obviousness at the 40% point between –a and a, that is, on the point of -0.2a. After the obviousness guidelines, assuming the examiner is still biased to the similar extent, the effective bar of obviousness would be at the 40% point between –b and b, on the point of -0.2b. Because the absolute value of a is larger than b, the obviousness guideline effectively raises the bar of obviousness, further reducing the number of obvious patents. The obviousness guideline would thus further reduce the social costs associated with bad patents, including deadweight loss and transaction costs.

Further, obviousness guideline may help the PTO to preserve patent incentives in an emerging field of technology. In the face of a fast emerging technology, the PTO may be concerned about examiners holding too low a standard of obviousness, over-granting patents and lowering patent quality. As a result, the PTO may want to resort to a conservative stance of categorically rejecting inventions in the field with a more rule-based doctrine, such as patentable subject matter.[173] The obviousness guidelines would alleviate this concern, making it is less likely that the PTO would adopt the categorical ban on a new field of technology, and preserving the incentives for research and invention in the new field.

Because the obviousness guideline would effectively raise the obviousness bar at patent prosecution, it would likely deter some patent applications, reducing the number of initial examinations at the first level of the patent funnel. The guideline would also provide a better tool for the examiner to reject an unworthy patent application, correcting the structural bias in the examination process to grant a patent, and further reducing the number of issued patents in the absence of the guidelines.

The guideline also likely saves patent prosecution costs for each application. The patent examination resembles a negotiation between the examiner and the patent applicant, a narrow margin of uncertainty leads to a narrow bargaining range between the parties, and thus will likely lower the negotiation cost for parties to reach an agreement.

C. Comparison with the AIA Proceeding

Besides nonlegislative rulemaking, the formal adjudications under the AIA proceedings provide another means by which the AIA can announce rules on

[173] Wasserman, *supra* note 52, at 420.

obviousness. Although AIA adjudications do not necessarily receive *Chevron* deference, these opinions are closely followed and analyzed by patent practitioners, and thus provide a good way for the PTO to convey information to the public at a time earlier than court litigation.[174] This section compares the costs and benefits of adopting obviousness rules through nonlegislative rulemaking with formal adjudications.[175] The actual obviousness rules will likely have similar contents as detailed in Part III, but these two approaches result in rules promulgated at different times, and with different initial costs.

The earlier an obviousness rule is promulgated, the more post-rule adjudications it will affect, and the more benefit it will confer on later adjudications.[176] For a Post Grant Review (PGR), the earliest of the AIA proceedings, it takes on average more than forty months from the initial filing of the patent to announce a rule on obviousness.[177] For nonlegislative rulemaking, an obviousness guideline could be promulgated in less than a year; a typical notice-and-comment rulemaking takes about six to twelve months, much earlier than the Post Grant Review.

On the other hand, nonlegislative rulemaking is likely to incur more costs than formal adjudication. An AIA proceeding likely will not incur much extra actual cost on the part of the PTO, other than opportunity cost, which exists

---

[174] The weight of these formal adjudication decisions is still unknown. Though the Federal Circuit has heard a few appeals from *inter partes* reviews, no opinion has been issued as of now. Arguments have been made that the formal adjudication decisions should receive Chevron deference, but even if they do not, they are likely to receive considerable respect from the court and the public for their persuasive power. The AIA proceedings are structurally similar to court litigations with discovery, oral arguments, and witness testimony. The parties are highly motivated to present their best arguments because of the preclusive effect of the proceedings. The presiding administrative patent judges are highly qualified patent attorneys who have technology backgrounds. They are probably better at understanding the technology-intense facts and the abstruse patent doctrines than normal juries and generalist trial judges.

[175] Though I use the term "rule" here, it only refers to the administrative law term of "rulemaking," and does not refer to the content of the rule. The content of the rule is detailed in Part III of this note, which could take various forms such as specific examples or a generalized rule. See *supra* Part III.

[176] This assumes that the rule does not affect private parties' research and development activities, which is likely true because the proposed obviousness rule only aims to narrow the margin of uncertainty, rather than change the standard of obviousness.

[177] The initial patent examination takes on average 27 months; a PGR petition must be filed within 9 months of issuance, 35 U.S.C. § 321; an institution decision should be made within 3 months of the preliminary response, 35 U.S.C. § 324, and the actual trial should be completed within a year from institution, 35 U.S.C. § 326(a)(11).

because the PTO has much less control over the rule made in an AIA proceeding than in a nonlegislative rulemaking. First, an AIA proceeding is initiated by a member of the public and litigated between private parties. The PTO has limited control over which AIA proceedings are initiated at what time, as well as the specific issues and arguments to be raised in the proceeding. A third party can challenge a wrongly granted patent under PGR within nine months of the grant of the patent,[178] or under an IPR, which does not have a time limit as in a PGR, but is still faster than a court litigation.[179] Second, the PTAB's adjudication function and the PTO's executive function are formally separated, and the executive function of the PTO has little sway in the decision-making of an APJ.

CONCLUSION

The core patentability standard of obviousness is a fact-specific, highly contextual inquiry that is notoriously hard to apply in a consistent way. The flexibility of the doctrine guarantees that each invention be evaluated on the whole scope of the facts in a case, but also creates problems of uncertainty and inconsistency, leading to type I and II errors. This is especially problematic at patent prosecution, when the PTO's structural bias for over granting patents combines with the indeterminateness of the obviousness doctrine, resulting in too many grants of obvious patents.

This note proposes that the PTO should promulgate obviousness guidelines under its authority to make nonlegislative rules. An obviousness guideline would include specific examples of inventions that are obvious, and potentially include generalized rules explaining why the inventions in the examples are obvious. Emerging fields of technology would be a good place to consider making such guidelines, because the obviousness doctrine are the most uncertain in these field.

An obviousness guideline would reduce the indefiniteness of the obviousness doctrine, reduce type I and type II errors, and provide a more stringent standard in counter of the structural bias in patent examination. Thus, it would reduce examination and litigation costs associated with an overflow of dubious patent applications and patents.

Despite the obvious benefits of promulgating such an obviousness guideline, the PTO has not done so in the past. This is likely because of the unique power

---

[178] 35 U.S.C. § 322 (2012).

[179] IPRs usually last a year; patent litigation in courts, on average, take 2 to 3 years. *Patent Litigation Cost*, INVENTIONSTATISTICS.COM, http://www.inventionstatistics.com/Patent_Litigation_Costs.html (last visited Mar. 27, 2016).

division between the agency and the courts in the patent area. The Federal Circuit takes a prominent role in making substantive patent policy and refuses to give the PTO *Chevron* deference. Accordingly, the PTO has perceived its own role as merely following the law made by Congress and the courts, rather than actively shaping the law. This is a shaky position because the PTO's examination duty necessarily requires the PTO to constantly make substantive patent law decisions. It is also a waste of the PTO's specialized expertise and knowledge that it can gain through its vital body of patent examination.

The obviousness guidelines provide a good starting point for the PTO to more proactively speak on substantive patent policy. Though the voice would be soft—without *Chevron* deference or binding power—it nonetheless enables the agency to claim a more active role in the U.S. patent regime in the post-AIA patent regime. The flexible nature of the obviousness doctrine ensures that the guidelines do not create substantive rights and obligations, but rather is tentative and flexible. The narrow scope of the guidelines would further make it easy for the PTO to draft the guidelines with adequate support and reasoning to make them persuasive to the courts.

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 5 SPRING 2016 NUMBER 2

HOW RINGS FIT INTO THE COPYRIGHT SCHEME: ASSESSING THEIR INTRINSIC UTILITARIAN FUNCTION

ADINE MITRANI*

This Note examines how rings fit into the copyright system as sculptural pieces not subject to the separability test under the useful articles doctrine. It focuses exclusive on rings, as they seemingly possess numerous functions; they have been used for mystical purposes, portable bank accounts, and as a signal of socially meaningful codes. Moreover, since jewelry designers consider functional features in the design process, should rings be treated as useful articles? After examining the purposes of the Copyright Act and conversing at length with designers, this Note concludes that rings are more of an art form. Although jewelry designers are limited by a finger's constraints, they employ a great deal of artistic creativity in expressing a message though the details of a ring.

* J.D. Candidate, New York University School of Law, 2016; B.A. International Relations, magna cum laude, University of Pennsylvania, 2012. The author would like to thank Professor Jeanne Fromer and the 2015-16 Editorial Board of the NYU Journal of Intellectual Property & Entertainment Law for their invaluable assistance throughout the editing process. The author would also like to acknowledge Courtney Crangi of Giles & Brother, Danielle Frankel Nemiroff of Phillips House, Tanya Gruenberg of Studio Grun, and Jeffry Aronsson of the Aronsson Group for speaking about their experiences with the jewelry design process and copyright protection.

INTRODUCTION

Jewelry is considered "one of the oldest art forms,"[1] dating further back than the history of painting and more vibrant than the history of sculpture.[2] Due to rampant imitation—or inspiration as some refer to it—in the fashion industry,[3] many jewelry designers seek to use the copyright system to protect their artistic designs.[4] In the past few years, retail stores such as Nasty Gal and Urban Outfitters

[1] NORMAN CHERRY, JEWELRY DESIGN AND DEVELOPMENT: FROM CONCEPT TO OBJECT 6 (Susan James et al. eds., 2013).

[2] Bruce Metcalf, *On the Nature of Jewelry*, JEWELRY AUSTRALIA NOW (1989), http://www.brucemetcalf.com/pages/essays/nature_jewelry.html.

[3] *See generally* Kal Raustiala & Christopher Jon Sprigman, *The Piracy Paradox: Innovation and Intellectual Property in Fashion Design*, 92 VA. L. REV. 1696-99, 1725–32 (2006).

[4] Interview with Courtney Crangi, CEO, Philip Crangi Jewelry and Giles & Brother, in New York City, N.Y. (Mar. 30, 2015) (explaining how she was quite upset when she saw a knockoff of a Giles & Brother signature necklace sold at a certain multinational retail store).

have been "slapped with lawsuits by jewelry designers who are crying foul over their work being mimicked without any credit and sold at fast fashion prices."[5]

Under the Copyright Act of 1976, many courts consider rings ornamental sculptures entitled to copyright protection as a pictorial, graphic or sculptural (PGS) work. With the objective of protecting artistic works and excluding functional designs from protection, the Copyright Act explicitly distinguishes useful articles. Specifically if a PGS qualifies as a useful article—defined as "having an intrinsic utilitarian function that is not merely to portray the appearance of the article or to convey information"[6]—then it is subject to a separability test. While courts have differed in their analysis in applying separability,[7] the test will essentially render the functional aspects of a PGS unprotectable.

Fashion designs are considered useful articles mostly unprotectable by Copyright system; their purpose to clothe people is utilitarian and the designs do not meet the Copyright Act's separability test.[8] Jewelry pieces are often lumped together with fashion articles and accessories, leading some to question their status as purely ornamental sculptures.

This Note focuses exclusively on rings as they seemingly possess numerous functional purposes as compared to other pieces of jewelry. For example, rings have been used for mystical and talismanic reasons. They are still used today as portable bank accounts, as the nature of small-scale jewelry allows wealth to be hidden easily and safely, to display socially meaningful codes, such as class rings and engagements rings, and to adorn one's hand.[9] Moreover, jewelry designers inevitably consider functional concerns when creating a ring; it is supposed to fit comfortably around one's finger. This raises the question that if a design is created with functional concerns, does it possess an intrinsic utilitarian function? Furthermore, there are diverging opinions in the federal district courts as to whether rings are considered sculptural works subject to a separability test.[10] While most courts hold that rings are purely artistic works, this Note seeks to examine

---

[5] Kathryn Dachille, *Bling It On: Copyright & The Rise of Jewelry Infringement Lawsuits*, CREATIVE ARTS ADVOCATE (2014), http://creativeartsadvocate.com/bling-it-on-copyright-the-rise-of-jewelry-infringement-lawsuits.

[6] 17 U.S.C. § 101 (2006).

[7] *See infra* Part I.B.2.

[8] Fashion designs do not meet the conceptual separability test because it is hard to distinguish their expressive and functional components. Hemphill & Suk, *The Law, Culture, and Economics of Fashion*, 61 STAN. L. REV. 1147, 1185 (2009).

[9] Cherry, *supra* note 1, at 6. *See also* Metcalf, *supra* note 2.

[10] *See infra* Part I.C.

why there has been confusion as to whether the useful articles doctrine applies specifically to rings.

Part I of this Note examines the history of jewelry in copyright, describing how courts have attempted to define intrinsic utilitarian function in light of the separability analysis. This section also compares case law that has expressly assessed rings in relation to the useful articles doctrine. Part II next considers rings as sculptural works not subject to a useful articles analysis. After briefly describing the evolution of jewelry in the 1960-80s, this section reinforces why rings qualify as ornamental sculpture for purposes of the Copyright Act. Part III considers rings as useful articles subject to a separability test, examining certain uses of rings as well as functional consideration in the design process. Lastly, this section analyzes critical reception in connection with rings, and explains how museum display is not a prerequisite for copyrightability. In the copyright spectrum with useful articles on one end and aesthetic objects on the other, this Note concludes that rings lie more on the artistic side. Rings do not possess an intrinsic utilitarian function for purposes of the Copyright Act, and therefore should not undergo a useful articles analysis.

I
HISTORICAL BACKGROUND OF JEWELRY IN THE COPYRIGHT ACT

Copyright law has included a variety of designs that are seemingly useful while simultaneously purporting to exclude any utilitarian products. Beginning in 1870, through a series of acts, Congress drew from piecemeal administrative and judicial formulations, in an attempt to reconcile the differences between protectable applied arts and unprotectable utilitarian designs. In 1949, Congress explicitly included "artistic jewelry" as within the scope of copyright protection. This section examines the history of jewelry in the copyright scheme, explains judicial attempts at defining intrinsic utilitarian function, and highlights district court cases that have expressly assessed rings in light of the useful articles doctrine.

A. Initial Encounters in Early Legislation and Case Law

Article 1, Section 8 of the Constitution authorizes federal legislation "[t]o promote the Progress of Science and useful Arts,"[11] but gives little guidance in defining the scope of the copyright system. The original Copyright Act of 1790 extended protection only to maps, charts, and books.[12] It was not until 1870 when

[11] U.S. CONST. art. I, § 8, cl. 8.
[12] *See* Act of May 31, 1790, ch. 15, 1 Stat. 124.

Congress explicitly extended copyright protection to three-dimensional objects: "painting, drawing, chromo, statute, statutary, and of models or designs intended to be perfected as works of the fine arts."[13] This statute purposefully used the term "fine art" in order to "maintain a respectable distance between copyright and useful articles."[14] The Copyright Act of 1909 eliminated this distinction, and seemingly allowed copyright coverage to apply to the designs of useful articles.[15] Specifically, the 1909 Act broadened the category of "fine arts" to include "[w]orks of art; models or designs for works of art."[16] One year later, however, the Copyright Office quickly corrected itself, amending the statute to expressly exclude "industrial arts utilitarian in purpose and character … even if artistically made or ornamented."[17]

With the advent of new useful articles in the early 20th century, such as television sets and new cosmetic products, it became harder to define the contours of industrial design. In 1917, the Copyright Act was reworded to cover "artistic drawings notwithstanding they may afterwards be utilized for articles of manufacture."[18] The Copyright Office promulgated a regulation in 1949 to expand its coverage and explicitly included "artistic jewelry."[19] Specifically, the regulation defined works of art as a class which "includes works of artistic craftsmanship, in so far as their form but not their mechanical or utilitarian aspects are concerned, such as *artistic jewelry*, enamels, glassware, and tapestries, as well as all works belonging to the fine arts, such as paintings, drawings and sculpture."[20] Thus, Congress established copyrightable categories with protection only covering the artistic elements of the designs.

In 1954, the Supreme Court decided *Mazer v. Stein*,[21] a seminal case for the useful articles doctrine, holding that copyright protection could be extended to

---

[13] Act of July 8, 1870, ch. 230, § 86, 16 Stat. 198, 212 (repealed 1916).

[14] Robert C. Denicola, *Applied Art and Industrial Design: A Suggested Approach to Copyright in Useful Articles*, 67 MINN. L. REV. 707, 711 (1983).

[15] *See* Act of March 4, 1909, 60th Cong., 2d Sess., 35 Stat. 1075. *See also*, Shira Perlmutter, *Conceptual Separability and Copyright in the Designs of Useful Articles*, 37 J. COPYRIGHT SOC'Y U.S.A. 339, 365 (1990).

[16] Act of March 4, 1909, 60th Cong., 2d Sess., 35 Stat. 1075.

[17] COPYRIGHT OFF., RULES AND REGULATIONS FOR THE REGISTRATION OF CLAIMS TO COPYRIGHT, Bulletin No. 15, at 8 (1910).

[18] Mazer v. Stein, 347 U.S. 201, 212 n.24 (1954) (quoting 37 C.F.R. § 201.4(7) (1917)).

[19] *Id*. at 212-213 (quoting 37 C.F.R. § 202.10(a) (1949)).

[20] *Id.* (emphasis added).

[21] *Id*. at 201.

sculptural figures that were used as bases for lamps.[22] *Mazer* attempted to distinguish artistic design, which qualifies for copyright protection, from ornamental features of useful articles, which belong to the design patent regime: "[t]he dichotomy of protection for the aesthetic is not beauty and utility but art for the copyright and the invention of original and ornamental design for design patents."[23] Based on the 1949 Regulation promulgated by the Copyright Office, *Mazer* supported the idea that artistic elements in useful articles would be entitled to Copyright protection as long as they remain physically separable from the utilitarian components.

Robert Denicola suggests that *Rosenthal v. Stein*[24] articulated a better approach to determine the Copyright Act's scope at the time.[25] Specifically, the Ninth Circuit in *Rosenthal* stated, "[a] thing is a work of art if it appears to be within the historical and ordinary conception of the term art."[26] Based on this definition, numerous cases upheld copyrights in jewelry, drawing on the historical conception of jewelry as a work of art.[27] Overall, before the Copyright Act of 1976, many courts considered artistic jewelry as a copyrightable category not subject to a separability test.

B. Further Defining Useful Articles

1. The Copyright Act of 1976

In § 102 of the Copyright Act of 1976, Congress approved these earlier precedents—specifically attempting to codify *Mazer*[28]—and enumerated eight categories of copyrightable subject matter. Section 102(a)(5) specifically included "pictorial, graphic and sculptural works" (PGS), thereby abandoning 1909's Act "works of art" categorization. While the new Act did not include specific

---

[22] *Id.* at 214.

[23] *Id.* at 218.

[24] 205 F.2d 633 (9th Cir. 1953).

[25] Denicola, *supra* note 14, at 17.

[26] 205 F.2d at 635.

[27] *See generally* Boucher v. Du Boyes, Inc., 253 F.2d 948 (2d Cir.), cert. denied, 357 U.S. 936 (1958); Dan Kasoff, Inc. v. Palmer Jewelry Mfg. Co., 171 F. Supp. 603, 606 (S.D.N.Y. 1959); Trifari, Krussman & Fishel, Inc. v. B. Steinberg-Kaslo Co., 144 F. Supp. 577 (S.D.N.Y. 1956); Trifari, Krussman & Fishel, Inc. v. Charel Co., 134 F. Supp. 551, 553 (S.D.N.Y. 1955).

[28] *See* H.R. REP. NO. 94-1476, at 54 [hereinafter HOUSE REPORT] ("In accordance with [*Mazer*] works of 'applied art' encompass all original pictorial, graphic, and sculptural works that are intended to be or have been embodied in useful articles, regardless of factors such as mass production, commercial exploitation, and the potential availability of design patent protection.").

examples, such as "artistic jewelry" from the 1949 regulation, many courts have held that jewelry is still a copyrightable sculptural work.[29] Furthermore, one court explained, "the explicit congressional adoption of the Copyright Office's definition indicates that jewelry remains within the scope of copyright protection."[30]

The 1976 Act also formulated the scope of useful articles: if a PGS meets the useful articles definition in § 101, "an article having an intrinsic utilitarian function that is not merely to portray the appearance of the article or to convey information,"[31] that article qualifies as useful. The design is not copyrightable unless its particular aesthetic elements satisfy the separability test. Section 101 defines the separability test as whether "such design incorporates pictorial, graphic, or sculptural features that can be *identified separately from*, and *are capable of existing independently of*, the utilitarian aspects of the article."[32] According to the legislative history, copyright protection for features of useful articles depended on whether the elements of the design are physically or conceptually separable from the utilitarian elements, such as the carving on the back of a chair or a floral relief design on silver flatware.[33] If the object is not separable, it does receive copyright protection. The following section identifies judicial attempts at defining intrinsic utilitarian function and determining a separability test. The last section examines how case law has treated rings in light of the useful articles doctrine.

2. Judicial Attempts at Defining "Intrinsic Utilitarian Function" and Separability

Many critics claim that the 1976 Act and its legislative history do not provide sufficient instructions to determine what the term "intrinsic utilitarian function" means.[34] While the House Report accompanying the 1976 Act identifies examples of "utilitarian articles such as textile fabrics, wallpaper, containers, and the like," it fails to define what constitutes an "intrinsic utilitarian function."[35] Courts have struggled to articulate an exact definition for intrinsic utilitarian function. *Gay Toys, Inc. v. Buddy L Corp.* attempted to distinguish the term from

---

[29] "It is undisputed that jewelry is included within the sculptural works classification of Section 102(a)(5)." Donald Bruce & Co. v. B. H. Multi Com Corp., 964 F. Supp. 265, 266 (N.D. Ill. 1997).

[30] Kieselstein-Cord v. Accessories by Pearl, Inc., 632 F.2d 989, 993 n.5 (2d Cir. 1980).

[31] 17 U.S.C. § 101 (2006).

[32] *Id.* (emphasis added).

[33] HOUSE REPORT, *supra* note 28, at 55 (1976).

[34] *See, e.g.*, Stephen Langs, *The Definitional Scope of an Intrinsic Utilitarian Function Under the 1976 Copyright Act: One Man's Use Is Another Man's Art*, 20 W. NEW ENG. L. REV. 143, 171-72 (1998).

[35] HOUSE REPORT, *supra* note 28, at 55 (1976).

other useful purposes, which may not rise to an intrinsic utilitarian function.[36] Other courts, such as *Brandir Int'l v. Cascade Pac. Lumber Co.* and *Poe v. Missing Persons*,[37] identified factors to determine usefulness. Moreover, courts have conveyed different approaches for establishing separability in a useful articles analysis. The Second Circuit in *Kieselstein-Cord v. Accessories by Pearl, Inc.* employed a consumer-based approach, focusing how the consumer uses the object to determine conceptual separability.[38] Five years later, the Second Circuit in *Carol Barnhart v. Economy Cover Corp.* established an object-based approach, assessing whether the function of the object drives the form to determine separability.[39] And in *Brandir*, the Second Circuit laid out a process-based approach, examining whether the form and function merge during the creation process to determine separability.[40] Below is a detailed analysis of how each case defines intrinsic utilitarian function in light of a separability analysis, and a brief overview of academic definitions.

In trying to determine whether belt buckles are copyrightable, Judge Oakes in *Kieselstein-Cord* did not delve into what, if any, intrinsic utilitarian function subsisted in the designs. Rather, he noted, "[t]he primary ornamental aspect of the [belt] buckles is conceptually separable from their subsidiary utilitarian function."[41] Throughout the opinion though, Judge Oakes failed to elaborate on what "subsidiary utilitarian function" means.[42] Ultimately, he upheld copyright protection for the etched metal belt buckles because some people wore them as jewelry, which the court determined was copyrightable subject matter.[43]

In *Carol Barnhart*, the Second Circuit looked at the function of the object to determine if four life-size polystyrene mannequins of human torsos were

---

[36] 703 F.2d 970, 973 (6th Cir. 1983) (noting that while a toy airplane may have other uses, its intrinsic purpose is to portray a real airplane).

[37] 834 F.2d 1142 (2d Cir. 1987) (examining the differences in design between the wire sculptural work and the ultimate bike rack, the utilitarian reasons in implementing the design changes, manufacturing concerns, advertising costs, and promotional or marketing objectives to determine whether the bike rack was a useful article); 745 F.2d 1238 (9th Cir. 1984) (looking at expert evidence, the designer's intent, testimony regarding industry practice in the art world and clothing trade, and marketing data to assess whether the bathing suit was a useful article).

[38] 632 F.2d 989, 993–94 (2d Cir. 1980).

[39] 773 F.2d 411, 419 (2d Cir. 1985).

[40] 834 F.2d at 1148.

[41] 632 F.2d at 994.

[42] Stephen Langs, *The Definitional Scope of an Intrinsic Utilitarian Function Under the 1976 Copyright Act: One Man's Use Is Another Man's Art*, 20 W. NEW ENG. L. REV. 143, 155-56 (1998)

[43] Kieselstein-Cord, 632 F.2d at 993–94.

protectable sculptural works under the Copyright Act.[44] In fact, Judge Newman's dissent expressly noted the "intrinsic" functional purpose of the mannequins: to "serv[e] as a means of displaying clothing and accessories to customers of retail stores."[45] After evaluating the legislative history and prior case law, the court held that the mannequins were not copyrightable because their function drove their form; mannequin surfaces are inextricably intertwined with the utilitarian purpose of displaying clothes.[46] Therefore, the court ruled that the mannequins were unprotectable as useful articles.[47]

Two years later, in *Brandir*, the Second Circuit looked at a number of different factors to determine whether the plaintiff's bike rack was useful: the differences in design between the wire sculptural work and the ultimate bike rack, the utilitarian reasons in implementing the design changes, manufacturing concerns, advertising costs, and promotional or marketing objectives.[48] To determine separability, the court mainly looked at the design process.[49] It held that even though the bike rack is worthy of admiration for its aesthetic qualities alone, utilitarian concerns during the creation process significantly influenced the design.[50] Specifically, the plaintiff expanded the undulating "sine-curve" of an artistic sculpture in order to accommodate it for bikes; accordingly, the court could not establish separability for purposes of the Copyright Act.[51]

In *Gay Toys*, the Sixth Circuit attempted to carve out a definitional difference between utilitarian function and intrinsic utilitarian function to determine whether toys, specifically model airplanes, are copyrightable.[52] Judge Brown stated that designs might have uses that go beyond portraying the appearance of the object or conveying information, but those uses do not necessarily constitute the intrinsic utilitarian function of the object.[53] He acknowledged that toys are designed for children to play with.[54] Yet in terms of the

---

[44] 773 F.2d at 412.

[45] *Id.* at 420. Ultimately, Judge Newman found the design features of the mannequins could constitute "conceptual separability." *Id.* at 426.

[46] *See id.* at 419. *See also* Shira Perlmutter, *Conceptual Separability and Copyright in the Designs of Useful Articles*, 37 J. COPYRIGHT SOC'Y U.S.A. 339, 365 (1990).

[47] 773 F.2d at 419.

[48] 834 F.2d 1142, 1146–48 (2d Cir. 1987).

[49] *Id.*

[50] *Id.*

[51] *Id.*

[52] 703 F.2d 970 (6th Cir. 1983).

[53] *Id.* at 973.

[54] *Id.*

Copyright Act, between unprotectable useful articles and protectable paintings, "[t]he function of toys is much more similar to that of works of art than it is to the 'intrinsic utilitarian function' of industrial products."[55] The court held that toys are copyrightable subject matter as a PGS, and are not subject to a useful articles analysis.[56]

A year later, in *Poe*, the Ninth Circuit reinforced *Gay Toys* definitional difference between usefulness and intrinsic utilitarian function.[57] Specifically, Judge Alarcon acknowledged that the swimsuit in question might have uses that go beyond portraying its own appearance, and remanded the case to the district court to determine whether such uses constitute an intrinsic utilitarian function.[58] The Ninth Circuit identified four factors that are relevant to the usefulness inquiry:

> (1) expert evidence may be offered concerning the usefulness of the article and whether any apparent functional aspects can be separated from the artistic aspects…(2) evidence of Poe's intent in designing the article may be relevant in determining whether it has a utilitarian function…(3) testimony concerning the custom and usage within the art world and the clothing trade concerning such objects also may be relevant;…and (4) the district court may also consider the admissibility of evidence as to Aquatint No. 5's marketability as a work of art.[59]

Similar to *Brandir*, these factors highlight the object's functionality in light of its aesthetic elements.

After analyzing relevant case law to determine the meaning of intrinsic utilitarian function, Professor Hick defines the term as objects that are designed with a specific function in mind (a factor mentioned in *Brandir* and *Poe*), such as a hammer that is designed to drive nails into a surface.[60] The extrinsic function, however, is derived from how consumers interact with the product—if one uses a hammer as a paperweight, that would constitute an extrinsic function of the hammer.[61] Furthermore, Thomas Byron explains how one might affix a coiled

---

[55] *Id.*
[56] *Id.* at 974.
[57] 745 F.2d 1238 (9th Cir. 1984).
[58] *Id.* at 1243.
[59] *Id.*
[60] Darren Hudson Hick, *Conceptual Problems of Conceptual Separability and the Non-Usefulness of the Useful Articles Distinction*, 57 J. COPYRIGHT SOC'Y U.S.A. 37, 53 (2010).
[61] *Id.*

extension cord to a wall or use a Van Gogh painting to swat flies.[62] Yet, displaying the extension cord as art, does not eradicate the primary purpose of the object; nor does using a famous artwork for a mundane task implicate that the art piece is less intrinsically artistic.[63] Therefore, he concludes, while many objects "may serve both aesthetic and utilitarian ends, different objects intrinsically serve these ends in varying degrees."[64]

In sum, courts and academics have treated utility and aesthetics as operating on a spectrum. At one end, there are objects that are purely functional, such as certain types of hardware. At the other end, there are inherently aesthetic works, such as a painting. In between, there are more questionable works, namely belt buckles and toys, which incorporate both aesthetic and utilitarian features in the design. In determining whether the design is subject to a separability test, the Copyright Act stipulates that it must possess an intrinsic utilitarian function, which courts have failed to adequately define. While *Kisselstein Cord* explicitly identifies jewelry as a protectable aesthetic category of the Copyright Act (not subject to useful articles analysis), the latter four cases seem to question the categorization of jewelry as a purely sculptural work.

C. Cases Assessing Ring Design Against the Useful Articles Doctrine

Based on courts' interpretations of the Copyright Act, it seems that rings are copyrightable subject matter as long as they meet requirements of authorship and originality, with the latter being the more difficult to prove.[65] While one district court held that rings are useful articles, two district courts explicitly rejected this argument and instead categorized rings as protectable ornamental sculptural pieces.

In *DBC of New York, Inc. v. Merit Diamond Corp.*,[66] the Southern District of New York took the position that rings are useful articles and proceeded with a

---

[62] Thomas M. Byron, *As Long as There's Another Way:* Pivot Point v. Charlene Products *as an Accidental Template for a Creativity-Driven Useful Articles Analysis*, 49 IDEA: THE INTELL. PROP. L. REV. 147, 181 (2008).

[63] *Id.*

[64] *Id.*

[65] *Compare* Weindling International, Corp. v. Kobi Katz, Inc., No. 00 Civ 2022, 2000 WL 1458788, *2 (S.D.N.Y. Sep. 29, 2000) (determining that combination of unoriginal elements in design of diamond "bridge" rings were original), *and* Diamond Direct, LLC v. Star Diamond Group, 116 F. Supp. 2d 525, 529 (S.D.N.Y. 2000) (considering the originality of diamond ring designs), *with* Vogue Ring Creations, Inc. v. Hardman, 410 F. Supp. 609, 611 (D.R.I. 1976) (holding that the ring design was "utterly devoid of any 'original creativity'" and therefore not protectable).

[66] 768 F. Supp. 414 (S.D.N.Y. 1991).

useful articles analysis.[67] The court explained how the rings' configuration—marquis stones flanked in a trillion ring setting—did not exist independently of its utilitarian counterparts.[68] The court did not include a reason why it categorized the diamond rings as useful articles; rather, the court, citing *Carol Barnhart*, reasoned, "the design of a useful article ... shall be considered a pictorial, graphic or sculptural work only if, and only to the extent that such design incorporates pictorial, graphic, or sculptural features that can be identified separately from, and are capable of existing independently of, the utilitarian aspects of the article."[69] After briefly conducting a separability analysis, the court held that the plaintiff's rings did not warrant copyright protection.[70]

Two district courts explicitly disagreed with the court's treatment of rings as a useful article in *DBC of New York*. Nearly a decade after, the Southern District in *Weindling Int'l, Corp. v. Kobi Katz, Inc.* stated that rings are "chiefly works of art, or more precisely ornamental sculptures, even if mass-produced."[71] After examining the combination of the rings' unoriginal elements—flaring supports, channel setting, triangle cut-outs, and sharp-edged apexes—it concluded that, in combination, the bridge ring merited copyright protection (specifically known as compilation protection).[72] Additionally, the Northern District Court of Illinois in *Donald Bruce & Co. v. B. H. Multi Com Corp.* stated that it did not agree with the defendant's reliance on the incorrect ruling in *DBC of New York*.[73] In determining whether plaintiff's Skalet Ring Line was copyrightable, the court flatly rejected the defendant's argument that rings can serve as useful articles saying, "[t]he [r]ing is purely ornamental, its sole purpose is to portray its appearance."[74] Accordingly, the court concluded that since the ring is not a useful article under § 101 of the Copyright Act, it therefore did not need to "determine whether the utilitarian aspects of the mount are separable from the sculptural elements."[75] After looking at the Skalet Ring Line's originality and assessing the validity of the copyright, the court dismissed defendant's motion for summary judgment.

---

[67] *See id.* at 416-417.
[68] *Id.* at 416.
[69] *Id.*
[70] *See id.* at 417.
[71] Weindling Int'l, Corp. v. Kobi Katz, Inc., No. 00 Civ. 2022, 2000 WL 1458788, at *4 (S.D.N.Y. Sep. 29, 2000).
[72] *Id.*
[73] 964 F. Supp. 265, 267 (N.D. Ill. 1997).
[74] *Id.*
[75] *Id.*

While the Copyright Act of 1976 did not expressly include the term "artistic jewelry," for the most part, rings are considered sculptural works. The absence of any reasoning behind the useful articles analysis for the diamond rings in *DBC of New York* further demonstrates that rings should be treated on the aesthetic end of the copyright spectrum.

II
CONSIDERING RINGS AS SCULPTURAL WORKS NOT SUBJECT TO A USEFUL ARTICLES ANALYSIS

This section examines how rings are considered works of art from jewelry designers' perspective and subsequently not subject to a useful articles analysis. After delineating a brief history on jewelry, this section will bolster what it means to be an "ornamental sculpture" for purposes of the Copyright Act, looking at three relevant factors in ring designs: artistic creativity, how jewelry designers view themselves and their creations, and how rings differ from clothing, a useful article.

A. A Brief History on Jewelry: The Rise of Artistic Expression

Jewelry is an old tradition steeped in the artistic world,[76] predating the history of sculpture and even painting.[77] Yet, rings can be considered useful in many senses. Rings have taken the forms of seals for legal documents, contracts, and international treaties, have been used for talismanic reasons and amuletic properties,[78] and are still used today as portable bank accounts and to display socially meaningful codes.[79]

Paul Greenhalgh delineates how the "design" category emerged from the "decorative arts." Specifically, in the 19th century, "[t]he decorative arts steadily congealed into a salon de refuse of genres that cohered only by virtue of their exclusion [from the category of fine arts]. Outside the fine arts, there was no fixed nomenclature or hierarchy. Variously—and interchangeably—known as the decorative, useful, industrial, applied or ornamental arts, they struggled to maintain a place in intellectual life …."[80] At the end of the nineteenth century, artists engaged in craft ethic differentiated themselves from those who produced large-scale manufacturing creations, which became known as design.[81] One of the

[76] Cherry, *supra* note 1, at 6.
[77] Metcalf, *supra* note 2.
[78] Cherry, *supra* note 1, at 6.
[79] Metcalf, *supra* note 2.
[80] Paul Greenhalgh, *The History of Craft, in* THE DESIGN HISTORY READER 329 (Grace Lees, Maffei & Rebecca Houze, eds., 2010).
[81] *Id.* at 333.

purposes of rings is to adorn,[82] and therefore rings can remain as a decorative art, distinguished from a fine art. Yet, because many jewelry designers mass-produce their creations, rings can also be considered designs.

Interestingly, though, in the height of the Arts and Crafts revolution from 1970-1980 in the United States, "jewelry had begun to take a new meaning in the art world."[83] This period has been compared to the "new painting" age that developed throughout Europe between 1880-1910 and subsequently in the United States from 1940-1960.[84] The late 1960s generated revolutionary jewelry students, especially potent in the art world, who challenged longstanding conventions and questioned the very notion of jewelry.[85] With the advent of new materials and metals, jewelry designers were tooled with new ways to create "individualistic art."[86] Specifically, jewelry designers employed non-precious metals and other soft metals as well as recycled materials, new plastics, and other refractory metals that could be colored in novel ways.[87] For instance, Crangi explained how she was one of the first designers to employ rugged brass in her creations when the company launched in 2001.[88] These innovative changes in jewelry design resulted in an invigorating and challenging debate whether jewelry was still "Art with a capital A"[89,90] Overall, jewelry has evolved to a forum for "conceptual exploration and personal expression."[91]

B. Considering Rings as Works of Art

In terms of the 1976 Copyright Act, many courts have articulated that rings are not useful articles because they are ornamental sculptures with a sole purpose

---

[82] HELEN W. DRUTT ENGLISH & PETER DORMER, JEWELRY OF OUR TIME: ART, ORNAMENT AND OBSESSION 12 (1995).

[83] Cherry, *supra* note 1, at 9.

[84] ENGLISH & DORMER, *supra* note 82, at 12.

[85] Cherry, *supra* note 1, at 9.

[86] ENGLISH & DORMER, *supra* note 82, at 12.

[87] Cherry, *supra* note 1, at 9.

[88] Interview with Courtney Crangi, *supra* note 4.

[89] Andy Warhol and others challenged the notion of what is art, namely "Art with a capital A." Many artists, philosophers, and writers began to take the view that art is how the viewer experienced the work, rather than the conventional rules prescribed to the creative piece: "if the artist offered a work of art and a viewer saw and understood it as art, then it must be art." Cherry, *supra* note 1, at 10.

[90] *Id.*

[91] Metcalf, *supra* note 2.

to convey appearance.[92] Moreover, artistic expression lies in how designers choose materials, how they work the materials to transform an object, and ultimately breathe new life into it.[93] Jewelry design is said to be the "highest level of craftsmanship and creativity—not blindly making [jewelry pieces] to preordained templates, but thinking through making, applying individual philosophies, personal intellect, active intuition, sensitivity, continuously inquiring and experimenting, and immense passion ...."[94] By including perspectives from jewelry designers, this section will further explain how a ring is an ornamental sculpture for purposes of the Copyright Act. Specifically, it will examine three relevant factors in ring design: artistic creativity and how the Copyright Act has aimed to protect that process; jewelry designers' perspectives and infusion of artistic concepts; and how rings differ from clothing, a useful article.

1. Artistic Creativity: Discovery in the Process

The history of the Copyright Act illustrates Congress's purpose in protecting artistic works. The Act has a rich history of explicitly including "fine art"[95] and "work[s] of art."[96] While the Act currently stipulates that PGSs qualify as a category of copyrightable subject matter, it further explains the phrase with references to art in the definitions section, saying PGS include: "works of fine, graphic, and applied *art*, photographs, prints and *art* reproductions …. Such works shall include works of *artistic* craftsmanship insofar as their form but not their mechanical or utilitarian aspects are concerned ...."[97] There are also policy concerns behind the Copyright Act that seek to incentivize "authors" to create artistic works. Many academics have supported the utilitarian justification for copyright protection: the Constitution authorized copyright legislation "to promote the Progress of Science and useful Arts,"[98] thereby suggesting an incentive rationale to encourage innovative activity through a system that promotes exclusive rights for the tangible results of creative efforts.[99] Accordingly, Professor

---

[92] 17 U.S.C. § 101. *See also* Weindling Int'l v. Kobi Katz Inc., No. 00 Civ. 2022, 2000 WL 1458788 at *4 (S.D.N.Y. Sep. 29, 2000); Donald Bruce & Co. v. B. H. Multi Com Corp., 964 F. Supp. 265, 267 (N.D. Ill. 1997).

[93] Cherry, *supra* note 1, at 121.

[94] *Id.*

[95] Act of July 8, 1870, ch. 230, § 86, 16 Stat. 198, 212 (repealed 1916).

[96] Act of March 4, 1909, 60th Cong., 2d Sess., 35 Stat. 1075 (repealed 1976).

[97] 17 U.S.C. § 101 (emphasis added).

[98] U.S. CONST. art. I, § 8, cl. 8.

[99] Denicola, *supra* note 14, at 722. *See* Mazer v. Stein, 347 U.S. 201, 219 (1954) ("The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to

Fromer argues that while the drafters of copyright legislation may not have had "artistic creativity" in mind, "[copyright]'s standard of originality for protectability, principally aligns with our understanding of how creativity proceeds and is valued in the artistic realm."[100] Specifically, the low bar for originality incentivizes authors to create works that are not "highly original," rather drawing on previous artistic works, which are typically in line with consumers' tastes.[101] Many ring designers employ artistic creativity, which further supports the notion that rings are sculptural works not subject to a useful articles analysis.

While scientific creativity seeks to find practical solutions for problems, artistic creativity focuses on the problem, posing questions to an audience to provoke debate and discussion.[102] The spirit of open-ended inquiry coincides with the Copyright Act's requirement of originality, as the author identifies a problem and fixes it in a work of art.[103] While jewelry is not considered a fine art in the strict sense, many jewelry designers see themselves as using artistic creativity in the process.[104] Norman Cherry interviewed seventeen jewelry designers, many who describe their creative process in artistic terms. For example, Simon Cottrell creates his designs in a structured form of improvisation, similar to how a jazz musician develops an initial theme and, through a complex combination of prior knowledge, experience, and intuition, eventually reaches a final destination.[105] Additionally, Ruudt Peters conveys that alchemy is a process in jewelry design that transmutes "prosaic materials into a visual poetry."[106]

Moreover, jewelry designers are actively engaged in a method referred to as "discovery of the problem," which involves both "deciding which artistic medium, materials, and represented objects will be used" and "harnessing experiences and themes for artistic expression."[107] Most jewelry artists do not rigidly follow a blueprint in the design process; they employ a number of different methodologies

---

advance public welfare through the talents of authors and inventors in 'Science and useful Arts."). *See also* Jeanne C. Fromer, *A Psychology of Intellectual Property*, 104 NW. U. L. REV. 1441, 1457-58 (2010).

[100] Fromer, *supra* note 99, at 1492.

[101] *Id.* at 1492, 1497.

[102] *Id.* at 1444. *See also* discussion *infra* Part II.B.2 (explaining how jewelry designers infuse of artistic concepts into their designs).

[103] Fromer, *supra* note 99, at 1444.

[104] *See generally* Cherry, *supra* note 1.

[105] *Id.* at 42.

[106] *Id.* at 92.

[107] Fromer, *supra* note 99, at 1467.

for their creations.[108] Gruenberg states that "[m]uch of what I do is instinctive … I don't consciously follow a step-by-step process to construct a ring."[109] When creating the ring template for 3D printing, she says that she designs, experiments with components, reacts, and then redesigns—like drawing through materials.[110] During the process new problems may arise and the artist is tempted to reframe that problem.[111] For instance, Peter Skubik designed a mercury ring by creating a mold, pouring liquid mercury into the mold and cooling the mercury in carbon dioxide snow to solidify it.[112] Once he removed the solidified ring from the mold, the ring melted a little and lost its form.[113] It created a dripping effect, calling attention to the malleable metal in relation to the sturdy mold.[114] Lastly, Courtney Crangi commented on her brother Phillip Crangi's designs saying that during college he noticed how "steel and gold love to live together;" he has since created a fine jewelry line that includes rings that combine the two materials.[115]

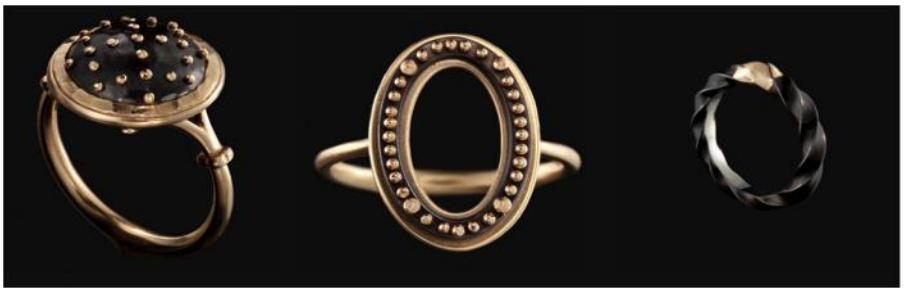


Figure 1: Phillip Crangi's Fine Jewelry Line[116]

Accordingly, copyright law rewards many artists and writers for articulating a particular emotion or subjective concept into a tangible work, rather than "only one problem solution receives the prize of copyright," as in the patent regime.[117]

---

[108] Cherry, *supra* note 1, at 121.

[109] Interview with Tanya Gruenberg, Creative Director/Jewelry Designer, Studio Grun, in New York City, N.Y. (Mar. 3, 2015).

[110] *Id.*

[111] Cherry, *supra* note 1, at 16.

[112] *Id.* at 106.

[113] *Id.*

[114] *Id.*

[115] Interview with Courtney Crangi, *supra* note 4; *see also* PHILLIP CRANGI DESIGNS, http://www.philipcrangi.com/collections/womens-rings (last visited May 19, 2015).

[116] PHILLIP CRANGI DESIGNS, *supra* note 115.

Overall, the Copyright Act's low originality threshold is designed to award the use of artistic creativity. Many jewelry designers describe their creation process in poetic terms and focus on the discovery of a problem, which mirrors the creative process of other artists, musicians, and writers. Thus, jewelry should be considered on the aesthetic side of the copyright spectrum not subject to a useful articles analysis as it incentivizes jewelry designers to engage in artistic creativity.

2. Designers' Perspectives

As the *Brandir* and *Poe* courts noted, one factor in determining usefulness is the designer's intent in creating the object. Specifically, Professor Hick explains that the intrinsic utilitarian function is mainly derived from how the designer purposed the object.[118] Accordingly, it is important to see how jewelry designers view their pieces to determine whether rings qualify as useful articles. To many designers and academics, jewelry is considered "one of the most vibrant, exciting and challenging contemporary art forms."[119] Many jewelry designers see themselves as artists.[120] In line with the 1976 Copyright Act's objectives, jewelry designers have compared their work to that of a sculptor.[121] Similar to sculptural objects, rings "consist of a physical object that has its own discrete existence."[122] In fact, Skubik views his rings in isolation from the body and sees them as art even when it is not worn.[123]

While there is a commercial component to jewelry design, Tanya Gruenberg, the creator of Studio Grun jewelry designs, states that she manages the production as a businesswoman, yet creates as an artist.[124] For sales, she has to think formulaically, keeping track of her cash flow and inventory.[125] As an artist, she sees her rings as a platform for investigating the interplay of a diverse range of mixed media—3D printing, plastic molds, and gemstones—to highlight a "unique degradation aesthetic while maintaining elegant accents."[126] Other jewelry designers infuse different art world concepts into their designs, such as abstraction,

---

[117] Fromer, *supra* note 99, at 1493.
[118] Hick, *supra* note 60, at 53.
[119] Cherry, *supra* note 1, at 10.
[120] Metcalf, *supra* note 2. *See* Cherry, *supra* note 1, at 13-17.
[121] *See e.g.*, *Designer Profile: Philip Crangi*, STYLE CASTER (2009) http://stylecaster.com/designer-profile-philip-crangi/ ("I see jewelry as small sculptural piece, so really for me, this is a sculptural pursuit.").
[122] Metcalf, *supra* note 2.
[123] Cherry, *supra* note 1, at 105.
[124] Interview with Tanya Gruenberg, *supra* note 109.
[125] *Id.*
[126] *Id.*

minimalism, and color schemes. Along with her husband, Emmy van Leersum creates pieces in an abstract manner that "rejected the craft tradition of the silversmith and strove instead to eradicate any trace of personal expression from their creations."[127] Additionally, Otto Künzli highlights social critique, frequently disparaging the pretentions of jewelry in his pieces.[128]

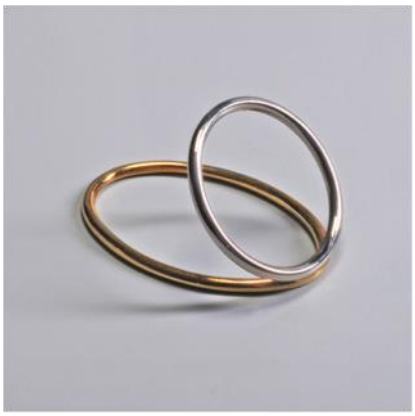


Figure 2: Emmy van Leersum – Broken Lines Ring[129]

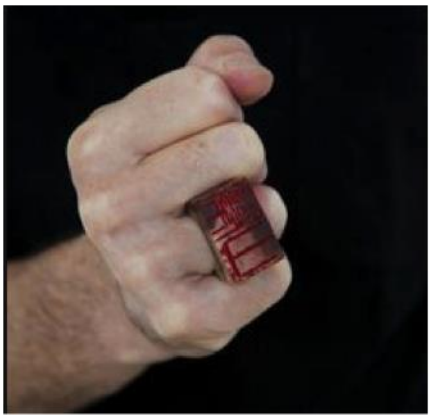


Figure 3: Otto Künzli – Seal Ring[130]

[127] CURRENT OBSESSION, *The Gijs + Emmy Show*, http://current-obsession.com/THE-GIJS-EMMY-SHOW (last visited March 19, 2016).

[128] DOMUS, *Otto Künzli Jewellery*, http://www.domusweb.it/en/news/2014/08/18/otto_kunzli_jewellery.html (last visited March 19, 2016).

[129] CHP JEWELRY, http://shop.chpjewelry.com/design-jewelry-chp22-broken-lines-ring-emmy-van-leersum-gold (last visited May 19, 2015).

[130] PINTEREST, https://www.pinterest.com/pin/468655904947185241/ (last visited May 19, 2015).

Lastly, Irene Neuwirth, a native from South California, explains that her inspiration is the ocean: "[i]ts purity, power and colors are all key elements at the origin of her designs."[131]

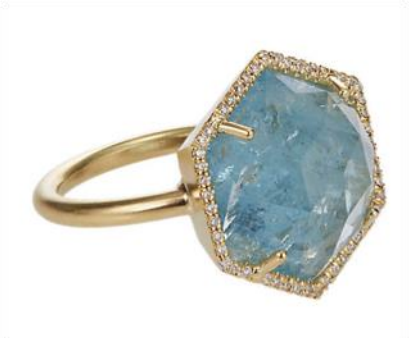


Figure 4: Irene Neuwirth – Gemstone Hexagonal Ring[132]

Overall, many jewelry designers see themselves as artists and incorporate many artistic themes into their pieces, thereby highlighting how jewelry is more of an art form for purposes of the Copyright Act.

3. Adornment Purposes of Rings: Distinct from Clothing

While other useful articles certainly incorporate artistic styles in their designs,[133] rings are inherently different from other industrial designs. As mentioned, the copyright system treats fashion pieces as useful articles,[134] and most jewelry designs as sculptural ornamental pieces.[135] However, many designers would submit that the nature of rings is innately tied to the human finger as rings

[131] IRENE NEUWIRTH JEWELRY, http://ireneneuwirth.com/ (last visited May 19, 2015).

[132] BARNEYS NEW YORK, http://www.barneys.com/irene-neuwirth-gemstone-hexagonal-ring-503107450.html (last visited May 19, 2015).

[133] *See e.g.*, MINIMALISSIMO MAGAZINE, http://minimalissimo.com/2009/08/humidifier/ (last visited May 19, 2015) (a minimalist humidifier); ARCHITECTURE, ART & DESIGN, http://www.architectureartdesigns.com/27-cool-furniture-ideas-inspired-by-pop-art/ (last visited May 19, 2015) (pop art chairs).

[134] *See* United States Copyright Office, Rules and Regulations for the Registration of Claims to Copyright 12(g) (1917) (listing "garments" among works that should not be registered).

[135] *See* Weindling Int'l v. Kobi Katz Inc., No. 00 Civ. 2022, 2000 WL 1458788, at *4 (S.D.N.Y. Sep. 29, 2000); Donald Bruce & Co. v. B. H. Multi Com Corp., 964 F. Supp. 265, 267 (N.D. Ill. 1997).

are made—and imagined—to be worn.[136] In this regard, rings are similar to garments in that the body is the site of the creative work. However, rings are remarkably different in the sense that they do not protect people from cold temperatures, which could be considered an intrinsic utilitarian function of clothing.[137] Additionally, clothing is used for modesty purposes in many cultures,[138] whereas rings could be said to have a complete opposite function—they mainly serve an adornment purpose: "it beautifies, within the value system of the local culture, and sometimes renders the wearer socially or sexually desirable."[139] Of course, fashion designs can accomplish the same objective, but the copyright system does not treat that as a primary objective of clothing.[140] Unlike clothing, which in most parts of the world you are legally required to wear, donning rings is a choice.[141] Many jewelry designers would submit that the primary purpose of rings is to decorate fingers and convey a specific style.[142]

Similar to clothing, though, jewelry designers must cater to trends, and in order "to be complete," their designs should be purchased and worn by others.[143] Many traditional artists are generally free from those "constraints of commerce,"[144] whereas many jewelry designers have to develop a business sense. Furthermore, English and Dormer argue that art "confers a status upon an object that is currently higher than and different from the status of craft or design" seen in jewelry.[145] However, the Copyright Act still considers many sculptural works as works of art, even if they are mass-produced or marketed for commercial purposes.[146] Judge

---

[136] Metcalf, *supra* note 2.

[137] *Id.*

[138] ART, DESIGN & VISUAL THINKING: AN INTERACTIVE TEXTBOOK, http://char.txa.cornell.edu/art/dress/dress.htm (last visited May 19, 2015).

[139] Metcalf, *supra* note 2.

[140] *See* Raustiala & Sprigman, *supra* note 3, at 1745–55.

[141] *Id.*

[142] ENGLISH & DORMER, *supra* note 82, at 20 ("Ornament and decoration are still important objectives in jewelry design.").

[143] *Id.* at 13.

[144] *Id.*

[145] *Id.* at 13-14.

[146] *See* HOUSE REPORT, *supra* note 28, at 54 ("In accordance with [*Mazer*] works of "applied art' encompass all original pictorial, graphic, and sculptural works that are intended to be or have been embodied in useful articles, regardless of factors such as mass production, commercial exploitation, and the potential availability of design patent protection."); *see also* Mazer v. Stein, 347 U.S. 201, 217-18 (1954) (We find nothing in the copyright statute to support the argument that the intended use or use in industry of an article eligible for copyright bars or invalidates its registration.").

Rakoff in *Weindling International*, points out "[a]rtistic design, after all, is at the very heart of the jewelry business, even in its crasser commercial forms."[147]

Furthermore, some designers see their rings as a more "intimate" art form between the creator and audience.[148] The weight of a ring combined with its "texture and size" on one's finger and size serves as a tactile "constant reminder of its presence."[149] Moreover, a ring, after being worn for years, "acquires a patina of age and even conforms itself to the shape of the finger."[150] While clothing and jewelry share similar elements, a ring's primary purpose is to adorn, which makes it more of an ornamental sculptural work for purposes of the Copyright Act. Additionally, the copyright system and its low threshold for originality is designed to afford tangible rights to the artistic creative efforts of designers. Since, many jewelry designers engage in artistic creativity (focusing on the problem) and infuse their designs with artistic concepts (expressing a particular problem), their designs should be protected as aesthetic works.

III
CONSIDERING RINGS AS USEFUL ARTICLES SUBJECT TO A SEPARABILITY TEST

As mentioned, jewelry is considered a sculptural work under the Copyright Act. The Copyright Act of 1976 further distinguishes between sculptural works and sculptural works that are useful articles with an "intrinsic utilitarian function," and affords protection to the latter as long as their artistic features "can be identified separately from, and are capable of existing independently of, the useful design."[151] Given that rings do possess some functional uses, this section assesses whether those uses constitute an "intrinsic utilitarian function" in terms of the Copyright Act. Both *Gay Toys* and *Poe* have explained that not all uses rise to the level of intrinsic utilitarian function.[152] Continuing, this section examines functional aspects underlying ring designs, such as those considered in *Brandir*, and assess whether that is an important factor in determining intrinsic utilitarian function. Lastly, the section analyzes critical reception of rings, and how the copyright system considers museum display.

---

[147] Weindling Int'l v. Kobi Katz Inc., No. 00 Civ. 2022, 2000 WL 1458788 at *4 (S.D.N.Y. Sep. 29, 2000).

[148] Metcalf, *supra* note 2.

[149] *Id.*

[150] *Id.*

[151] 17 U.S.C. § 101 (defining a "useful article" as "an article having an intrinsic utilitarian function that is not merely to portray the appearance of the article or to convey information").

[152] *See supra* Part I.B.2.

A. The Purposes of Rings: Examining their Intrinsic Functionality

A ring can possess a number of functional purposes; it can be used as an amuletic power, a portable bank account, a signifier of status, and an allurement to one's hands. After examining each of these uses, this section will conclude that some of these uses are more extrinsic functions, not rising to the level of intrinsic utilitarian function for purposes of the Copyright Act.

Some people buy rings for the talismanic power of the gemstones. In line with popular culture, the American Gem Society designates a different gemstone that contains a unique mystic meaning to each calendar month. For instance, Garnet, the birthstone of January, is "known to promote romantic love, passion, sensuality, and intimacy."[153] Accordingly, Jewelry by December 1967, a designer on Etsy, lists a Mozambique Garnet Ring and explains "[g]arnet is said to be the stone of romantic love and passion, enhancing sensuality, sexuality, and intimacy."[154]

Figure 5: Jewelry by December 1967 – Mozambique Garnet Ring[155]

Additionally, the Mystical Maven, another seller on Etsy, presents a Golden Garnet Ring, which she describes as a ring that will allure the opposite sex if the other person touches the ring while the owner is wearing it.[156]

---

[153] AMERICAN GEM SOCIETY, *The Healing Power of Gemstones*, http://www.americangemsociety.org/healing-gemstones (last visited May 19, 2015).

[154] ETSY, https://www.etsy.com/listing/233352938/mozambique-garnet-ring-sterling-silver?ga_order=most_relevant&ga_search_type=handmade&ga_view_type=gallery&ga_search_query=garnet%20ring%20love&ref=sr_gallery_42 (last visited May 19, 2015).

[155] *Id.*

[156] ETSY, *Golden Garnet Ring*, https://www.etsy.com/listing/232005178/golden-garnet-ring-love-spellinfused?ga_order=most_relevant&ga_search_type=handmade&ga_view_type=gallery&ga_search_query=garnet%20ring%20love&ref=sr_gallery_2 (last visited May 19, 2015).

Moreover, certain cultures attach mystic powers to certain gemstones, such as the Chinese who consider the rich luster of jade to be very lucky.[157] In illustrating the talismanic forces of the jade stone, Melissa Chang conveys, "[w]hen [my dad] first started wearing [the jade ring], he told me that the deeper the green was, the more good luck it brought. He always wore it to Vegas and as far as I know, he did pretty well on the craps table."[158] While these amuletic powers of rings can qualify as a function, it would be quite difficult to prove actual, intrinsic utilitarian function from anecdotes. The question also remains: to what extent should mystic powers qualify as utilitarian? Indeed, the *Gay Toys* court stated "[t]he intention of Congress was to exclude from copyright protection industrial products such as automobiles, food processors, and television sets."[159] Mystical purposes do not seem to be in line with Copyright Act's intention of excluding useful industrial objects.

Rings are still used today as portable bank accounts because the nature of small-scale jewelry allows wealth to be hidden easily and safely.[160] While this may not be the traditional use, this use constitutes an extrinsic function of a ring, in the same manner as when a person uses a hammer as a paperweight.[161] Jewelry designers generally do not design a ring so that it can be quickly converted into cash. As mentioned in Part II.B.ii, jewelry designers see themselves as artists and incorporate many artistic elements into their designs. Moreover, various paintings could be said to serve as portable bank accounts in the form of investment art, yet, that feature does not prevent courts from categorizing paintings as purely aesthetic and entitled to copyright protection.

Rings can also be used to display socially meaningful codes.[162] For the most part, though, rings have shifted from conveying class identity, to being more

---

[157] CULTURAL CHINA, *A Brief Introduction of Chinese Jade-Culture*, http://arts.cultural-china.com/en/32Arts4663.html (last visited May 19, 2015).

[158] JADE GALLERY HAWAII, *The Power of Jade*, http://www.jadegalleryhawaii.com/power-of-jade.htm (last visited Feb. 3, 2016).

[159] Gay Toys, Inc. v. Buddy L Corp., 703 F.2d 970, 973 (6th Cir. 1983). These categories are subject to patent protection, which requires "novelty," a more demanding standard. *See generally* Joseph Scott Miller, *Hoisting Originality*, 31 CARDOZO L. REV. 451, 465-67 (2009).

[160] Metcalf, *supra* note 2.

[161] *See supra* Part I.B.2.

[162] *See* Metcalf, *supra* note 2. *See also* Kristen Booker, *What Happens When the World Doesn't Understand Your Hair*, MARIE CLAIRE (Mar. 4, 2015, 4:23 PM), http://www.marieclaire.com/beauty/news/a13591/what-happens-when-the-world-doesnt-understand-your-hair/ (commenting on the status of a David Yurman ring and how NY subway goers find it unlikely for an African-American woman with curly hair in a full twist out to own a real one).

stylistic choices that can distinguish the wearer or merge the wearer with a particular style.[163] According to the Copyright Act's definition of a useful article, though, social identity markers can be analyzed as merely conveying information, and therefore jewelry should not be considered useful in that sense. Moreover, a number of objects convey social status and identity, such as a handbag or a car, but that does not necessarily speak to their intrinsic utilitarian function; a handbag is known to hold items and a car is known as a method of transportation.

Lastly, donning rings certainly draws attention to one's hands, "appeal[ing] to potential mates."[164] As *Gay Toys* alluded to, many purely aesthetic objects can have some uses.[165] Art pieces do not only serve informational or decorative purposes; they can also arouse passions, offer escape, and serve as a forum of dialogue and contemplation.[166] Drawing attention to one's hands can serve as an aesthetic experience, heightening the artistic essence of the ring. Furthermore, that visual experience bolsters the idea that rings fall more on the aesthetic side of the copyright spectrum. After all, even though an art piece can have functional characteristics derived from emotive power or historical meaning, such characteristics do not make it a useful article under copyright law.[167]

In light of the legislative history behind the Copyright Act and case law, these aforementioned uses do not rise to the level of intrinsic utilitarian function. The legal doctrine of functionality should not be broadened to include extrinsic functions or further aesthetic functions, as that can render even an art piece a useful article.

B. Functional Considerations Behind Ring Design

While jewelry designers take into account a number of considerations in designing a ring, they inevitably take function into account. Alice Sprintzen, a jewelry designer who wrote an instructive book on basic jewelry techniques, emphasized the importance in accounting for functional concerns in the design

---

[163] Possibly with the exception of engagement rings, which can still serve as class identifiers. *See The History of Jewelry: Why do we Wear Jewelry?*, SAY WHY DO I (Sept. 3, 2011), http://www.saywhydoi.com/the-history-of-jewelry-why-do-we-wear-jewelry/.

[164] *See id.*

[165] Gay Toys, Inc. v. Buddy L Corp, 703 F.2d 970, 973 (6th Cir. 1983).

[166] Hick, *supra* note 60, at 53. "Alternatively, several theorists have argued that the function of art is to bring about the 'aesthetic experience,' to produce 'aesthetic contemplation,' or to yield 'aesthetic satisfaction." *Id*.

[167] Melissa M. Mathis, *Function, Nonfunction, and Monumental Works of Architecture: An Interpretive Lens in Copyright Law*, 22 CARDOZO L. REV. 595, 621 (2001).

process.[168] Specifically, a designer should consider the weight of the item, flexibility of the material, any protrusions on a piece that might catch clothing, and external circumstances.[169] In the section about wax casting a ring, Sprintzen instructs designers to slightly enlarge wide band rings to allow for finger swelling in hotter seasons.[170] Danielle Frankel Nemiroff, one of the co-designers for Phillips House fine jewelry said, "functionality and comfort are not the leading concerns in the design process, but definitely are important" in the design process.[171] Nemiroff's design philosophy is to create "pretty pieces" that one can wear comfortably; she mentioned that she will not create a ring with a sharp spike that can potentially hurt a young child or damage an evening gown.[172]

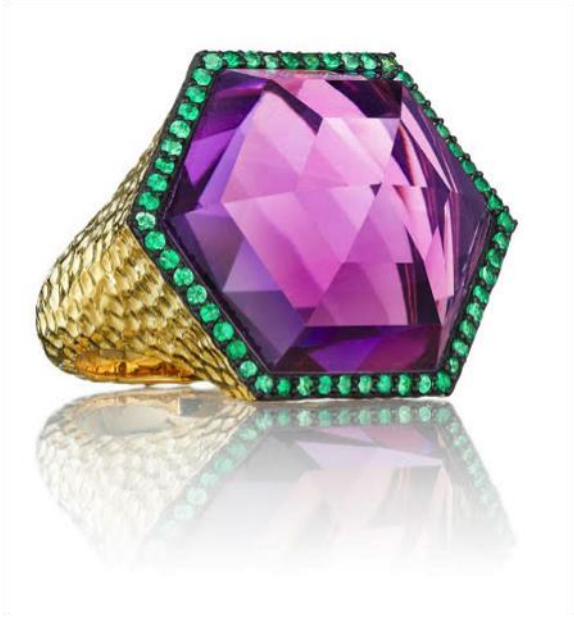


Figure 6: Phillips House – No. 3 Hexagon Ring[173]

[168] *See* ALICE SPRINTZEN, JEWELRY, BASIC TECHNIQUES AND DESIGN, 6 (1980).
[169] *Id.*
[170] *Id.* at 74.
[171] Interview with Danielle Frankel Nemiroff, Co-Designer, Phillips House, in N.Y.C., N.Y. (Apr. 2. 2015).
[172] *Id.*
[173] *No. 3 Hexagon Large Ring*, THE SHOP BY PHILLIPS HOUSE, http://shop.phillipshouse.com/collections/rings/products/yg-am-no-3-large-hex-ring (last visited May 19, 2015).

While Gruenberg sees herself as artist, she says she also takes ergonomics and negative space into account when designing a ring, especially for her King Crown Ring that goes above and below the knuckle.[174]

Figure 7: Studio Grun – King Crown Ring[175]

Specifically, she tries to create lighter, more comfortable rings.[176] Additionally, Jennie Kwon, another jewelry designer, stated that she does not want women to be burdened in removing her pieces when washing their hands; rather she creates delicate designs that will not be cumbersome to the wearer.[177]

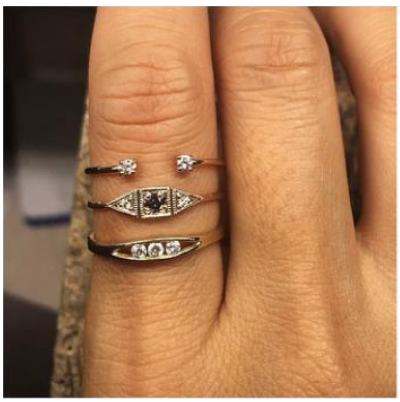


Figure 8: Jennie Kwon – Black Diamond Mini Deco Point Rings[178]

---

[174] Interview with Tanya Gruenberg, *supra* note 109. Studio Grun cuffs, however, are designed heavier to signify luxury. *Id.*

[175] *King Crown Ring*, STUDIO GRUN, http://www.studiogrin.com/shop/companion-ring-turquoise (last visited Feb. 18, 2016).

[176] Interview with Tanya Gruenberg, *supra* note 109.

[177] *See* Maia Adams, *Fine Jeweller Jennie Kwon*, ADORN JEWELLERY BLOG (May 9, 2014), http://www.adorn-london.com/profile/jennie-kwon/.

[178] *Black Diamond Mint Deco Point Ring*, JENNIE KWON DESIGNS, http://jenniekwondesigns.com/products/black-diamond-mini-deco-point-ring (last visited May 19, 2015).

Yet, does accounting for functional concerns during the design process bestow an intrinsic utilitarian function on the object? While the *Brandir* court analyzed the creation process as a main factor in determining separability,[179] it does not necessarily indicate that the object has an intrinsic utilitarian function. Due to the nature of the ribbon bike rack, which was derived from a wire sculpture, the *Brandir* court collapsed the useful article inquiry with the separability analysis in trying to establish copyrightable subject matter.[180] Thus, whether the design process was heavily influenced by utilitarian concerns should be more of a question for separability, especially when considering the primary purpose of the object. In essence, ring designers are creating for a finger, their canvas, and are naturally limited by that medium's constraints. Rings are worn on the human finger, which imposes functional considerations of weight, height, and scale, but there still remains considerable room for creativity and expression.[181]

C. Critical Reception: Does It Matter?

Art collectors, art museums and critics in the art market do not typically think of jewelry as a fine art, such as a painting or a sculpture.[182] Specifically, English and Dormer maintain that the Museum of Modern Art or the Venice Biennale do not feature jewelry prominently in their collections, and it rarely appears in art periodicals such as *Art Forum*.[183] The Museum of Art and Design, however, recently presented a "stunning array of extravagant fashion jewelry" from June 2013–April 2014.[184] Additionally, from November 2013–March 2014, the Metropolitan Museum of Art had an exhibit titled Jewels by JAR (Joel A. Rosenthal) with bedazzled, vibrant jewelry pieces using the pavé technique.[185] It was the first retrospective at the Museum tailored to a contemporary artist of gems.[186] Yet, these examples perhaps highlight jewelry as more of a craft, rather than an art. Indeed, the Museum of Art and Design, formerly known as the Museum of Contemporary Crafts, has exhibits that feature other useful articles

---

[179] *See* Brandir Int'l v. Cascade Pac. Lumber Co., 834 F.2d 1142, 1148 (2d Cir. 1987).

[180] *See id.* at 1143, 1148.

[181] Metcalf, *supra* note 2.

[182] *See* ENGLISH & DORMER, *supra* note 82, at 14

[183] *Id.*

[184] MUSEUM OF ART & DESIGN, *Fashion Jewelry: The Collection of Barbara Berger*, http://madmuseum.org/exhibition/fashion-jewelry (last visited May 19, 2015).

[185] THE METROPOLITAN MUSEUM OF ART, *Jewels by JAR*, http://www.metmuseum.org/about-the-museum/press-room/exhibitions/2013/jewels-by-jar (last visited May 19, 2015).

[186] *Id.*

such as "Ralph Pucci: The Art of the Mannequin."[187] Olaf Skoogfors, one of the main jewelry designers involved in the evolution of the abstractionist, formal approach, said, "I consider myself to be an artists as well as a craftsperson. The same efforts that go into painting or sculpture go into my jewelry. If this medium is a lesser art, than I am a lesser artist."[188] Even though jewelry can be considered a craft, it still falls more on the aesthetic side of the copyright system, which is perhaps bolstered by a recent expanded notion of art in the museum world.

Furthermore, whether jewelry is featured in a museum is not a primary concern for purposes of the Copyright Act. While the *Kieselstein* court noted that two of the belt buckles at issue in the case were placed on display in the Metropolitan Museum of Art in ruling that the buckles were protected by copyright,[189] museum placement merely serves as an additional argument for copyright protection rather than a prerequisite. Judge Rakoff in *Weindling International* commented that it was highly doubtful that the bridge ring in question would appear in the Met, but noted how "the law of copyright protects the modest creations of the humble versifier who churns out greeting cards as much as it does the thrilling inventions of a poet laureate."[190] After all, the Supreme Court in *Feist Publications v. Rural Telephone Co.* defined the originality requirement as merely "some minimal degree of creativity…some creative spark, 'no matter how crude, humble, or obvious' it might be."[191] Because mass-produced jewelry will inevitably be designed in non-original ways, it still may warrant copyright protection "if the creative spark behind a commercial jewelry design is more like a flickering match than a bolt a lightning."[192] Therefore, originality remains a primary prerequisite for aesthetic works.

In sum, rings should not be considered useful articles subject to a separability test. While rings certainly have uses, these uses do not qualify as an

---

[187] MUSEUM OF ART & DESIGN, *Ralph Pucci, The Art of the Mannequin* http://madmuseum.org/exhibition/ralph-pucci (last visited May 19, 2015); *cf.* Carol Barnhart v. Economy Cover Corp., 773 F.2d 411 (2d Cir. 1985) (holding a mannequin as a useful article).

[188] ENGLISH & DORMER, *supra* note 82, at 20.

[189] Kieselstein-Cord v. Accessories by Pearl, Inc., 632 F.2d 989, 991 (2d Cir. 1980).

[190] Weindling Int'l v. Kobi Katz Inc., No. 00 Civ. 2022, 2000 WL 1458788 at *4 (S.D.N.Y. Sep. 29, 2000) (probably referring to *Roth Greeting Cards v. United Card Co.*, 429 F.2d 1106 (9th Cir. 1970), which granted Copyright protection to greeting card that paired a simple, trite drawing on the outside with the phrase "I love you" on the inside.)

[191] 499 U.S. 340, 345 (1991). Although the Court acknowledged that "the requisite level of creativity is extremely low" and that "even a small amount will suffice," it concluded that white-page telephone directories fail to meet the Copyright Act's originality requirement. *See id.*

[192] Weindling Int'l, Corp., 2000 WL 1458788, at *4.

intrinsic utilitarian function in light of case law and the legislative history behind the Copyright Act. While jewelry designers generally account for functional considerations in the design process, these considerations serve more as a factor for determining separability. While the designer is inevitably going to run into practical issues by designing a ring for a finger, a great deal of artistic expression can exist within those confines. Lastly, museum placement is an added benefit in asserting copyrightability, and rings can still be an aesthetic work despite some historical notions that it served as a craft.

CONCLUSION

Ring design is an art in many senses: many jewelry designers use artistic creativity in making these ornamental sculptural pieces that leave an intimate relationship between the maker and bearer. At the same time, though, rings have been considered useful, especially when they were used as seals or to signify a class. Given the new wave of "individualistic art" during the height of the Arts and Craft Revolution, some of the old purposes of rings were diminished. As inventiveness has become a prominent objective in jewelry design, functionality is often neglected.[193] With the advent of new metals available during the 1960–80s, jewelry designers began to use a number of different materials and tools, employing new methods to create "individualistic art."[194] Considering themselves as artists, jewelry designers seldom create "straight 'social jewelry'" with familiar meanings such as class rings or wedding/engagement rings; "[s]tripped of familiar codes and functions, jewelry has become a vehicle for purely artistic issues"[195] in line with the Copyright Act's objectives.

After examining the purposes behind the Copyright Act and conversing at length with designers, I argue that jewelry rings lie more on the aesthetic side of the copyright spectrum. Since they do not possess an intrinsic utilitarian function, they should not be subject to a useful articles analysis. The process of designing rings certainly takes functional features into account,[196] and could seem confusing against the backdrop of the useful articles doctrine. Yet, courts have pointed out that the primary purpose of rings is to adorn. Designers employ artistic creativity in creating rings, which the Copyright Act is designed to protect. Also, designers see

---

[193] *See* Metcalf, *supra* note 2.

[194] ENGLISH & DORMER, *supra* note 82, at 12.

[195] *See* Metcalf, *supra* note 2.

[196] *Cf.* Carol Barnhart v. Economy Cover Corp., 773 F.2d 411 (2d Cir. 1985) (deciding a mannequin was a useful article because the function was driving the form). *See also* Brandir Int'l v. Cascade Pac. Lumber Co., 834 F.2d 1142 (2d Cir. 1987) (holding a bike rack uncopyrightable because form and function merged during the design process).

themselves and their creations as art, which further indicates that rings do not contain an intrinsic utilitarian function. Categorizing the functions of a ring as intrinsically utilitarian would unduly expand the functionality doctrine and could render works of art as useful articles. While jewelry designers consider how a ring fits comfortably around a finger, they incorporate artistic elements to convey a personal story through the details. In articulating a particular expression in a tangible form, jewelry designers should be entitled to robust copyright protection for their rings—specifically, as sculptural works not subject to a useful articles analysis.

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 5 SPRING 2016 NUMBER 2

😀 © 📕: EMOJI AS LANGUAGE AND THEIR PLACE OUTSIDE AMERICAN COPYRIGHT LAW

RACHEL SCALL*

As emoji become more ubiquitous in society, users are learning to express themselves through these symbols. Copyright protection of emoji would hamper this growing area of free expression. This note argues that, given the ways in which emoji are used in American culture, they should not receive copyright protection, in order to encourage the use of emoji as an "accessory" to language. Emoji do not readily fit under U.S. copyright protection and their maintenance would be best left to private organizations. This structure would allow people to use emoji freely, in order to develop common meanings for symbols among emoji users and thereby maximize their communicative and expressive functions.

* J.D. Candidate, New York University School of Law, 2016. The author would like to thank the Editorial Board of the Journal of Intellectual Property & Entertainment Law, participants of the 2016 Innovation Policy Colloquium, and Professor Jeanne Fromer.

INTRODUCTION

In the mid-1990s, as email and internet communication were gaining ground in his native Japan, Shigetaka Kurita noticed a problem with these new means of communication. Traditionally, the Japanese people would communicate via long personal letters filled with lengthy phrases and greetings meant to convey emotions that were not necessarily found in the dictionary definitions of the words on the page.[1] Email involved much shorter and quicker communication. As a result, people left lengthy expressions of emotion off the page.[2] Suddenly it was not clear whether a given word in an email was "a kind of warm, soft 'I understand' or a 'yeah, I get it' kind of cool, negative feeling."[3]

Kurita recognized that online communications were likely to remain short and terse in comparison to Japan's traditionally long written letters. As such, he sought to find a new, shorter way to express the connotations of a traditional writer's written word. Drawing from street signs, Chinese characters, and symbols used in manga comics,[4] Kurita developed a series of symbols that represent emotions and other abstract ideas.[5] The symbols, which began life as a system of

[1] Jeff Blagdon, *How Emoji Conquered the World*, VERGE (Mar. 4, 2013, 11:46 AM), http://www.theverge.com/2013/3/4/3966140/how-emoji-conquered-the-world.

[2] *Id.*

[3] *Id.*

[4] Manga is a style of Japanese comics. Anime is the animated version of manga. *See Manga and Anime*, JAPAN GUIDE, http://www.japan-guide.com/e/e2070.html (last visited Feb. 22, 2016).

[5] *Id.*; *see also* Mayumi Negishi, *Meet Shigetaka Kurita, the Father of Emoji*, WALL ST. J. (Mar. 26, 2014, 5:36 AM), http://blogs.wsj.com/japanrealtime/2014/03/26/meet-shigetaka-kurita-the-father-of-emoji/.

176 12-pixel by 12-pixel glyphs,[6] eventually evolved into more than 1,000 symbols now known as emoji.[7]

Kurita "never expected emoji to translate abroad,"[8] but emoji became a mainstay of American culture after Apple included an emoji keyboard with its iPhone iOS 2.2 update in 2011.[9] Since then, emoji have worked their way into many aspects of online communication. Emoji are used in private communications, such as text messages and emails, and public communications, such as Twitter and blog posts.[10] Some users have truly taken their emoji usage to the next level; in 2009, Fred Benson founded a Kickstarter campaign to translate Herman Melville's Moby Dick into an all-emoji version, titled Emoji Dick.[11]

In their short life, emoji have had little contact with the American legal system. In the criminal trial context, courts have begun to admit evidence that includes emoji in the context of a text message or online posting.[12] Judge Katherine Forest, presiding over a trial concerning the online black-market website Silk Road, instructed the jury to take note of any emoji included in any document and

---

[6] Blagdon, *supra* note 1.

[7] Adam Sternbergh, *Smile, You're Speaking Emoji*, N.Y. MAG (Nov. 16, 2014, 9:00 PM), http://nymag.com/daily/intelligencer/2014/11/emojis-rapid-evolution.html; Amit Chowdhry, *Apple Releases iOS 8.3 to the Public, It Has New Emoji*, FORBES (Apr. 8, 2015, 1:17 PM), http://www.forbes.com/sites/amitchowdhry/2015/04/08/apple-releases-ios-8-3-to-the-public-its-the-update-with-the-new-emojis/. This paper will only analyze the emoji based on code provided by the Unicode Consortium. Different companies have begun developing their own "branded emoticons," but these glyphs function by downloading an app that provides a branded emoticon keyboard. *See generally* Kristina Monllos, *Here's Why Your Favorite Brands Are Making Their Own Emoji*, ADWEEK (Mar. 9, 2015, 9:15 PM), http://www.adweek.com/news/advertising-branding/here-s-why-your-favorite-brands-are-making-their-own-emoticons-163325. This is different from emoji, which come standard on a variety of devices and function cross-platform as in-line text, rather than inserted pictures.

[8] Negishi, *supra* note 5.

[9] Blagdon, *supra* note 1.

[10] *See, e.g.*, EMOJITRACKER, http://www.emojitracker.com/ (last visited May 14, 2015) (tracking real-time emoji usage on Twitter); EMOJINALYSIS, http://emojinalysis.tumblr.com/ (last visited May 14, 2015) (blogging about the psychology behind the "recently used" emoji on people's cell phones).

[11] Erin Allen, *A Whale of an Acquisition*, LIBR. CONGRESS (Feb. 22, 2013), http://blogs.loc.gov/loc/2013/02/a-whale-of-an-acquisition/. Today, a copy of Emoji Dick resides in the Library of Congress.

[12] Eli Hager, *Is an Emoji Worth 1,000 Words?*, MARSHALL PROJECT (Feb. 2, 2015, 3:34 PM), https://www.themarshallproject.org/2015/02/02/is-an-emoji-worth-1-000-words.

to consider the emoji part of any document submitted to evidence.[13] In the context of intellectual property law, however, little has been said about how emoji fit into the American system of intellectual property protections.

Given that emoji are, by their nature, digital code-based pictures, it is possible that if emoji were to receive intellectual property protection, such protection could fit into the copyright system. Copyright protects "original works of authorship fixed in any tangible medium of expression."[14] Facts and ideas themselves are not protected.[15] One form of copyrightable works is "pictorial, graphic, and sculptural works," to the extent that the works are aesthetic and not functional.[16] Emoji are two-dimensional representations that could plausibly fall under the pictorial, graphic, and sculptural category. Even if emoji do fall within the domain of copyright protection, however, public policy may dictate that they would be better left to the public domain.

This paper argues that, given the ways in which emoji are used in American culture, they should not receive copyright protection and should be left to the public domain. Copyright law should treat emoji more like an evolving language than intellectual property belonging to a person or entity. Americans already use emoji in their communications, and copyright's constitutional purpose of promoting the arts and sciences[17] would be best achieved by encouraging the use of emoji as an "accessory" to language. Free use of emoji as part of the American lexicon will promote communication, thus promoting learning and free expression.[18] Part I of this paper will discuss how emoji function. This section will explain how emoji work from a technical standpoint, as well as how they have become part of today's social and communicative framework. Part II will analyze how emoji could fit into current U.S. copyright law under the protections for pictorial, graphic, and sculptural works. This section will also discuss the impact such protection would have on the incentives to create emoji and will also touch on other intellectual property regimes that may be applied to them. Part III will explain why the public would most benefit from categorically excluding emoji from copyright protection, and how the growth of emoji should be governed outside copyright protection.

---

[13] Benjamin Weiser, *At Silk Road Trial, Lawyers Fight to Include Evidence They Call Vital: Emoji*, N.Y. TIMES (Jan. 28, 2015), http://www.nytimes.com/2015/01/29/nyregion/trial-silk-road-online-black-market-debating-emojis.html?_r=0.

[14] 17 U.S.C. § 102.

[15] Feist Publ'n, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991).

[16] 17 U.S.C. § 101.

[17] U.S. Const. art I, § 8, cl. 8.

[18] Golan v. Holder, 132 S.Ct. 873, 901 (2012) (Breyer, J., dissenting).

I
EMOJI'S TECHNICAL ELEMENTS AND THEIR COMMUNICATIVE FUNCTION

This section will explore how emoji function and how people use them. Part A will cover the technology and business of emoji and how they appear on users' devices. Part B will explore how emoji are developing as language as people begin to work them into various forms of electronic communication.

A. How Emoji Work from a Technical Standpoint

Shigetaka Kurita's first set of emoji, created in the mid-1990s,[19] were a feature on a pager marketed to teenagers.[20] A somewhat uniform set of emoji did not emerge until 2008, and the emoji most Americans are familiar with only emerged when Apple included the characters in a 2011 iOS update.[21]

Emoji may appear to be simply a series of pictographs of people (👯), places (🏭), and things (🎥), but, in essence, each emoji is a unique piece of computer code.[22] When Japanese technology companies first began to incorporate emoji in mobile technologies, such as pagers and cell phones, different companies used different codes to represent the same emoji symbol, and sometimes the same code to represent different symbols.[23] This coding problem was not unique to emoji: as different companies around the world entered the realm of computing, different methods emerged for coding symbols used in virtually all languages.[24] Different forms of coding for the same symbols – emoji or otherwise – led to interoperability between computing platforms.[25]

Enter the Unicode Standard. The goal of Unicode is to provide a "unique number for every character, no matter what the platform, no matter what the program, no matter what the language."[26] The first Unicode standard debuted in

---

[19] Blagdon, *supra* note 1.

[20] Jessica Bennet, *The Emoji Have Won the Battle of Words*, N.Y. TIMES (July 25, 2014), http://www.nytimes.com/2014/07/27/fashion/emoji-have-won-the-battle-of-words.html?_r=0.

[21] *Id.*

[22] Amy Weiss-Meyer, *A Peek Inside the Non-Profit Consortium That Makes Emoji Possible*, NEW REPUBLIC (June 27, 2014), http://www.newrepublic.com/article/118421/emoji-made-possible-non-profit-consortium.

[23] *Id.*

[24] *What is Unicode?*, UNICODE, http://www.unicode.org/standard/WhatIsUnicode.html (last visited May 14, 2015).

[25] *Id.*

[26] *Id.*

1992.[27] All modern Internet browsers and most leading operating systems support Unicode.[28] The Unicode Consortium manages the Unicode Standard through its role as a non-profit "founded to develop, extend and promote use of the Unicode Standard."[29] The Unicode Consortium has been working to standardize various characters for Internet use for more than 20 years.[30] Unicode's "work is ubiquitous to the point of being invisible";[31] the Unicode Standard governs every character that people read or type on electronic devices.[32]

Given that emoji emerged as a set of characters that only exist via technology,[33] it is not surprising that it should fall to the Unicode Consortium to manage the code that allows users to communicate via emoji.[34] The Unicode Consortium gives each emoji symbol a code and a name, such as "U+1F36D LOLLIPOP."[35] The name only generally describes the character, while the code instructs the computer what, specifically to pull up as text.[36]

The Unicode Consortium determines the code that makes the emoji appear on a user's screen, but it does not design the actual emoji seen by the user.[37] The emoji symbols are similar to typefaces, and are designed by each technology company that chooses to incorporate emoji in its product.[38] For example, U+1F49B YELLOW HEART is designed by apple to look like 💛, but is designed by Android to look like 💛.[39] There are other emoji displays that are simpler and resemble a traditional Dingbats font.[40] The Unicode Consortium provides the

---

[27] *Chronology*, UNICODE, http://www.unicode.org/history/versionone.html (last visited February 21, 2016).

[28] *Id.*

[29] *Id.*

[30] Weiss-Meyer, *supra* note 22.

[31] *Id.*

[32] *Id.*

[33] Blagdon, *supra* note 1.

[34] *Emoji and Dingbats*, UNICODE, http://www.unicode.org/faq/emoji_dingbats.html (last visited May 14, 2015).

[35] *Id.*

[36] *Id.*; *see generally Miscellaneous Symbols and Pictographs*, UNICODE, http://www.unicode.org/charts/PDF/U1F300.pdf (last visited May 14, 2015).

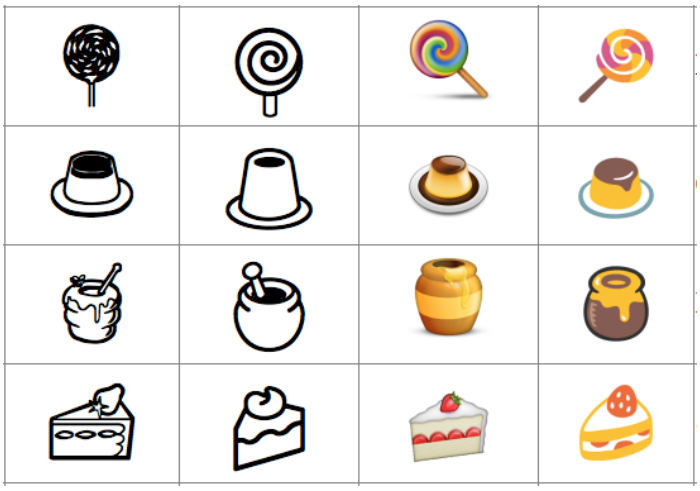
[37] Weiss-Meyer, *supra* note 22.

[38] *Id.*

[39] John-Michael Bond, *You May Be Accidentally Sending Friends a Hairy Heart Emoji*, ENGADGET (Apr. 30, 2014, 7:00 PM), http://www.engadget.com/2014/04/30/you-may-be-accidentally-sending-friends-a-hairy-heart-emoji.

[40] *Emoji and Dingbats*, *supra* note 34. Dingbats is a font made up of symbols, rather than alphanumerical characters.

following chart, which compares examples of different emoji displays used by different companies:



The chart features, from top to bottom, four different displays for U+1F36D LOLLIPOP, U+1F36E CUSTARD, U+1F36F HONEY POT, and U+1F370 SHORTCAKE.[41] Technology companies are free to display each piece of code as they choose, and the Unicode Standard names are provided as suggestions for how a given piece of code should appear to the user.[42]

To better understand the role of different players in the functionality of emoji, it is helpful to analyze a problem that has plagued American emoji users since Apple first popularized emoji: the dearth of racial diversity displayed by the "human" characters. Apple's original emoji character set featured more than 30 representations of humans, as well as various hand gestures and body parts (such as ears and noses), all of which were Caucasian.[43] Arguably only 2 or 3 emoji were not Caucasian, and not one was black.[44] Apple agreed there must be more emoji diversity, but claimed its hands were tied by the code provided by the Unicode Standard.[45] However, this statement appears to be an example of Apple dodging

---

[41] *Id.* The third column depicts Apple's interpretation of emoji, while the fourth column depicts Android's display.

[42] *Id.*

[43] *See generally* Ben Reid, *iOS 8.3, OS X 10.10.3 Adds New Emojis, Here's What They Look Like*, REDMOND PIE (Feb. 24, 2015), http://www.redmondpie.com/ios-8.3-os-x-10.10.3-adds-new-emojis-heres-what-they-look-like.

[44] *Id.*

[45] Joey Parker, *What Does Apple Think About the Lack of Diversity in Emojis? We Have Their Response.*, MTV (Mar. 25, 2014, 1:59 PM), https://web.archive.org/web/20140327033829/http://act.mtv.com/posts/apple-responds-to-lack-of-diversity-in-emojis/.

the diversity issue because the Unicode Standard does not control the appearance of the emoji or require any racial or ethnic manifestation.[46] Although it is clear that Unicode controls the emoji code, and companies like Apple control the way they look, the Unicode Consortium worked to find ways to partner with the companies designing emoji in order to provide more diversity.[47] Apple's release of iOS 8.3 in April 2015 finally gave users six skin-tone options to choose from for the majority of "people emoji," but the emoji representing families and couples are only available in a yellow, non-human skin tone.[48]

Because emoji code makes it possible to include images in-line with text, they occupy a new and unique way of communicating. Although users could previously share a photo with one another, or draw each other pictures, the versatility of emoji mixed with text gives them the potential to develop as part of language, or even as their own form of language. Now, people can replace words with emoji where they feel the emoji will convey more emotion simply a typed word.

B. Emoji as an Element of Language Expanding Communication

Emoji may have been intended to clarify connotations in brief online communications,[49] but today they play a larger role in digital communications.[50] Emoji can be used "as punctuation [excited face], as emphasis [sob], as a replacement for [several] words ("Can't wait for [palm trees] [sun] [swim]!") or to replace words altogether."[51] Emojitracker, an online database of real-time emoji use on Twitter, updates so quickly that it opens with an epilepsy warning.[52] In fact, according to Emojitracker's data, "people are averaging 250 to 350 emoji tweets a second."[53] This calculation does not even account for emoji used in text messages, email, "gchat,"[54] and other platforms.[55]

Emoji have been referred to as "an optional written language,"[56] "a foreign language,"[57] and "digital hieroglyphics that, in many cases, can substitute for

---

[46] *Emoji and Dingbats*, *supra* note 34.
[47] *Id.*
[48] *See* Reid, *supra* note 43.
[49] Blagdon, *supra* note 1.
[50] Bennet, *supra* note 20.
[51] *Id.*
[52] EMOJITRACKER, *supra* note 10.
[53] Bennet, *supra* note 20.
[54] Gchat is a colloquial term for Google's Gmail instant messenger system.
[55] *See* Bennet, *supra* note 20.
[56] Sternbergh, *supra* note 7.

lettered language."[58] Linguist Ben Zimmer has said that although emoji are not yet a "full-fledged language," they do "seem to have fascinating combinatorial possibilities. Any sort of symbolic system . . . used for communication[] is going to develop dialects."[59] Zimmer's statement is telling: emoji have been used for a number of purposes, but seem most effective when used in the context of an already existing full-fledged language. For example, Emoji Dick, an all-emoji translation of Moby Dick,[60] does not have quite the level of elegance and readability as the original, in part due to the fact that there are more than 1,000,000 words in the English language,[61] but only 1,000 or so emoji.[62] Emoji can, however, approve (👍), express emotions (😄 😓 😵), describe what you want for dinner (🍕), and add to written works and conversations in a variety of ways. In fact, linguist Tyler Schnoebelen has found that emoji have begun to develop their own grammar of sorts.[63] For example, emoji tend to appear at the end of messages.[64]

Emoji's growth as a language depends on the organic use and development of meaning that has been characteristic of its use thus far. Intellectual property protections, however, could stymie that growth by taking emoji out of the public domain. If emoji creators own copyrights in their creations, users may lose some of their freedom to transform emoji into an increasingly useful form of communication.

II
FAILURE TO FIT EMOJI INTO THE AMERICAN COPYRIGHT STRUCTURE

It is clear that emoji can potentially be classified in multiple ways: as a series of pictures, as typefaces, as computer code or as a component of language. How emoji fit into U.S. copyright protections depends on how they are classified.

---

[57] *See* Damon Darlin, *America Needs Its Own Emojis*, N.Y. TIMES (Mar. 7, 2015), http://www.nytimes.com/2015/03/08/opinion/sunday/turn-emojis-red-white-and-blue.html.

[58] Ruth Reader, *The Emoji Is the Future of Texting on the Apple Watch*, VENTUREBEAT (Mar. 9, 2015, 3:46 PM), http://venturebeat.com/2015/03/09/the-emoji-is-the-future-of-texting-on-the-apple-watch.

[59] Darlin, *supra* note 57.

[60] Allen, *supra* note 11.

[61] *Number of Words in the English Language*, GLOBAL LANGUAGE MONITOR (Jan. 1, 2014), http://www.languagemonitor.com/number-of-words/number-of-words-in-the-english-language-1008879/.

[62] *See* Sternbergh, *supra* note 7; Chowdhry, *supra* note 7.

[63] Katy Steinmetz, *Here Are Rules of Using Emoji You Didn't Know You Were Following,* TIME (July 17, 2014), http://time.com/2993508/emoji-rules-tweets/.

[64] *Id.*

Part A of this section will analyze copyright protection for pictorial, graphic, and sculptural works and how emoji – either individually or as a set – could fit into this category of copyright protection. Next, Part B will look at how copyright law would treat emoji when used in combination and will touch on how trademark law may govern combinations of emoji. Part C will briefly discuss possible copyright protection for emoji code. Part D will discuss the standard-essential patent model of intellectual property protection and the consequences of applying a compulsory licensing scheme to emoji.

A. Pictorial, Graphic, and Sculptural Works: Individual Emoji vs. Full Set of Characters

Under the Copyright Act of 1976, U.S. copyright law protects "original works of authorship fixed in any tangible medium of expression," including "pictorial, graphic, and sculptural works" ("PGS" works).[65] PGS works include two and three-dimensional artistic representations, including commercial art, fine art, and anything in between.[66] These works are protected for their artistic merit, but are not protected insofar as they may be useful or functional.[67]

Emoji could plausibly gain copyright protection as PGS works, given their nature as two-dimensional pictures. To be eligible for this protection, emoji would first have to be "original," or a work "independently created by the author" that possesses "at least some minimal degree of creativity."[68] Originality is not a very difficult standard to meet: "a work may be original even though it closely resembles other works so long as the similarity is fortuitous, not the result of copying."[69] An original work must simply "possess some creative spark, 'no matter how crude, humble or obvious' it might be."[70]

To determine whether emoji fulfill copyright's originality requirement, it is important to determine whether the idea behind a given emoji and the expression of said emoji are separable. For example, the "grinning face emoji," also known as the "smiley face emoji" or "happy face emoji," looks like 😀.[71] The idea of a smiley face is not protectable by copyright because ideas themselves do not fall

---

[65] 17 U.S.C. § 102.
[66] *Id.* at § 101.
[67] *Id.*
[68] *Feist*, 499 U.S. at 345.
[69] *Id.*
[70] *Id.* (citing 1 MELVILLE B. NIMMER & DAVID NIMMER, COPYRIGHT § 1.08[C][1] (1990)).
[71] *Grinning Face*, EMOJIPEDIA, http://emojipedia.org/grinning-face/ (last visited May 15, 2015). Emojipedia is a website that catalogues emoji with their picture and common descriptive names.

within copyright's scope of protections.[72] Smiley faces, as a category, represent "a larger private preserve than Congress intended to be set aside in the public market" under the monopoly created by copyright.[73] But the grinning face emoji is only one example of the many ways to express a smiley face. In fact, the emoji set alone contains many examples of smiley faces: 😀, 😁, 😃. The different smiley face emoji are expression of the idea of a smiley face, and copyright protection may be applied to expression of an idea, so long as the expression is an original creation.[74] Given that smiley faces may be expressed in many ways, and assuming that emoji were not copied from already-existing smiley faces, the grinning face emoji likely fulfills copyright's originality requirement.[75]

There is, however, an important exception to copyright's originality rule. When an idea is "very narrow, so that the topic necessarily requires if not only one form of expression, at best only a limited number . . . the subject matter would be appropriated by permitting the copyrighting of its expression."[76] This "merger" of expression with idea makes a work uncopyrightable and therefore, "when merger occurs, identical copying is permitted."[77] This means that if there is only one way, or very few ways, to create an image of a particular idea, that image will not receive copyright protection.[78] To give copyright protection to an image that captures the heart of an idea would be to ignore the idea-expression distinction and to give a copyright monopoly over something that Congress judges to belong in the public domain.[79] For example, there may be only a small number of ways to create an icon of a basketball or an American flag. If this is true, it is possible that the basketball emoji, 🏀, and American flag emoji, 🇺🇸, are not copyrightable because the idea behind those emoji merges with the images used. The more generic, and less creative, an emoji is, the more likely it will merge with the idea it represents and therefore be uncopyrightable.

---

[72] *Feist*, 499 U.S. at 345.

[73] Herbert Rosenthal Jewelry Corp. v. Kalpakian, 446 F.2d 738, 742 (9th Cir. 1971)

[74] *Id*. at 740-41.

[75] *Id.* at 741. Rather than perform this analysis for every emoji, this paper will assume each emoji passes muster under this originality test, while acknowledging that a court may find otherwise.

[76] Morrissey v. Proctor & Gamble Co., 379 F.2d 675, 678-79 (1st Cir. 1967) (internal citations omitted).

[77] Lotus Development Corp. v. Borland Int'l, Inc., 49 F.3d 807, n. 9 (1st Cir. 1995).

[78] *Id*.

[79] *Morrissey*, 379 F.2d at 678-79.

Different providers of emoji "fonts" have interpreted the Unicode emoji code in order to come up with different, and original, displays of emoji.[80] Where an emoji is more than an "indispensable, or at least standard" manner of displaying a given idea, it is likely copyrightable (assuming it was not copied from a preexisting work).[81] However, where an emoji is either the only way, or one of very few ways to express an idea, it is likely uncopyrightable and part of the public domain. Given that many emoji do not appear to be generic images, or duplicate images representing the same ideas,[82] this analysis assumes that the majority of individual emoji are not subject to the merger doctrine and are therefore copyrightable.

Next, in order to gain copyright protection, emoji must be "fixed in any tangible medium of expression."[83] "A work is 'fixed' in a tangible medium of expression when its embodiment in a copy . . . is sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for a period of more than transitory duration."[84] Digital displays of original images are considered "fixed" because they can be perceived by the user.[85] It does not matter that the user affects the display by choosing the emoji she wishes to use – the display is still considered "permanent or stable" for the purpose of fixation.[86] Assuming that emoji, as individual images, meet the copyright requirements of originality and fixation, it appears that they are eligible for copyright protection as PGS works.

Taken as a collection, it is not as clear that emoji are protectable under current copyright law. As a collection, emoji bear a close resemblance to a typeface. A typeface is "a design of an alphabet and other typographical symbols placed on devices" used in connection with printing, traditionally, and digital displays.[87] Congress had an opportunity to consider copyright protection of typefaces while preparing the Copyright Act of 1976, but Congress decided against

---

[80] *See supra* Part I.A.

[81] Hoehling v. Universal City Studios, Inc., 618 F.2d 972, 979 (1980).

[82] For example, Apple's iOS 8.3 emoji keyboard contains 13 emoji representing trains or trams, three emoji representing CDs, 4 mailbox emoji, 14 emoji hearts, and many other duplicative emoji.

[83] 17 U.S.C. § 102.

[84] *Id.* at § 101.

[85] Williams Elec., Inc. v. Artic Int'l, Inc., 685 F.2d 870, 874 (1982).

[86] *Id.*

[87] Eltra Corp. v. Ringer, 579 F.2d 294, 296 (1978). *See also* Adobe Sys. Inc. v. S. Software Inc., 1998 WL 104303, at *4 (N.D. Cal. Feb. 2, 1998). A font is the computer program used to "generate human readable typeface designs on computer screens, printers, and other devices. Monotype Imaging, Inc. v. Bitstream, Inc., 376 F.Supp. 2d 877, 882 (N.D. Ill. 2005).

protecting typefaces under copyright law.[88] This decision, to deny copyright protection to typefaces, was upheld by the Fourth Circuit in *Eltra Corp. v. Ringer*.[89] There, the court stated that a "typeface is an industrial design in which the design cannot exist independently and separately as a work of art." Due to this holding, and lack of protection under the Copyright Act, typeface has never received copyright protection.[90]

Legal commentators still debate whether typefaces are truly industrial design or are actually works of art deserving copyright protection.[91] Because typefaces are necessary for humans to communicate via writing, giving typeface creators lengthy copyright monopolies over their designs would not make sense because it could limit means of communication, thereby stymying free expression. Furthermore, copyright law cannot provide a monopoly on the limited ways to legibly express a letter.[92] It is possible that the creative aspects of a typeface simply cannot be separated from its usefulness. Given these arguments, denying typefaces copyright protection makes sense.

Emoji may fall into the category of "other typographical symbols," much like the ITC Zapf Dingbats typeface.[93] ITC Zapf Dingbats qualifies as a set of copyrighted symbols, but is not a copyrighted typeface.[94] Therefore, it is possible that emoji are protected as individual "copyrighted symbols," but as a system represent an uncopyrightable typeface. Excluding emoji from copyright protection, as a typeface, would create a protection regime for emoji that would favor users over creators, thus allowing emoji to develop as a language, rather than as a marketable good. Without copyright protection, users are free to include emoji in text messages, websites, and other written digital works and communications without the threat of a copyright suit. This system of non-protection promotes free

---

[88] Leonard Storch Enter., Inc. v. Mergenthaler Linotype Co., 1980 WL 1175, at *4 (E.D.N.Y. Aug. 8, 1980).

[89] *Eltra Corp.*, 579 F.2d 294.

[90] *Id.* at 298.

[91] *See generally* Phillip W. Snyder, *Typeface Design After the Desktop Revolution: A New Case for Legal Protection*, 16 COLUM.-VLA J.L. & ARTS 97 (1991) (describing the legal status of typefaces in the digital age and possible intellectual property protections for digital typefaces); Emily N. Evans, *Fonts, Typefaces, and IP Protection: Getting to Just Right*, 21 J. INTELL. PROP. L. 307, 337-39 (2014) (discussing how typefaces may fulfill requirements of copyright law, but should not fall under copyright protection).

[92] Evans, *supra* note 91, at 337.

[93] ITC Zapf Dingbats is a typeface consisting of pictorial symbols rather than alphanumeric symbols. When the word "emoji" is typed in ITC Zapf Dingbats, it appears as "□□□□□"

[94] Monotype Corp. PLC v. Int'l Typeface Corp., 43 F.3d 443, n.3 (9th Cir. 1994).

speech via emoji, thus allowing the greatest number of people to participate in helping them evolve as a language or component of language.

Of course, it is necessary that companies have the incentives to create emoji fonts, so that users have access to emoji images and not only the underlying code. Without images, emoji are not useful for communication. It is unlikely that a lack of copyright protection for emoji images would disincentivize creators, given the environment in which emoji are currently created. Electronic goods and software manufacturers develop emoji as a feature of a given smart phone or other digital platform. So long as consumers demand emoji on their digital devices, manufacturers will have incentives to include emoji typefaces on those devices. It is not necessary that many emoji typefaces be created, because the fewer the typefaces, the more homogenous emoji will be, and the easier it will be for different people to understand the characters.

Although people may prefer one style or expression of Unicode's emoji code to another style or expression, communication of emoji is dependent upon uniform display to promote understanding. For example, what appears to an Apple user as the hair flick emoji (💁) may not be recognizable on an Android device, where the same emoji appears like this: 💁.[95] A study of popular anthropomorphic emoji found that people do not interpret emoji the same way and that "[c]ommunicating across platforms . . . adds additional potential for misconstrual."[96] Having one standard set of emoji, rather than encouraging people or companies to create various expressions of emoji, would help eliminate this communication problem.[97] Additionally, having a standard display would prevent various creators of emoji display from interpreting Unicode emoji code as completely different displays, as with the Apple yellow heart emoji and the Android 4.4 hairy heart emoji. A uniform system of emoji images would prevent different dialects from developing – a valid public policy goal that would avoid people being split into dialect groups simply based on their chosen electronic or mobile device provider. Otherwise, people with iPhones would end up developing one emoji dialect, while people with Android would essentially develop a separate language. Copyrighting emoji code would essentially place barriers between people simply because of their electronics preferences.

---

[95] EMOJIPEDIA, http://emojipedia.org/information-desk-person/ (last visited Feb. 22, 2016).

[96] Hannah Miller, et al., *"Blissfully happy" or "ready to fight": Varying Interpretations of Emoji*, ICWSM-16 at 8-9 (forthcoming May 2016), available at http://grouplens.org/site-content/uploads/Emoji_Interpretation.pdf.

[97] *See id.* at 9.

B. Intellectual Property Protection May Be Available for Emoji in Combination

When people use emoji to communicate, they may express their ideas using a grouping or combination of emoji. For example, you could ask someone if they want to grab pizza and a movie (🍕 ➕ 🎥 ❓), or signal disbelief with the phrase "holy crap" (🙏 💩).[98] Looking at the emoji set as a means of communication, each emoji represents a word or phrase. Groupings of emoji therefore make up phrases or sentences, but would likely not be considered creative works for the purpose of copyright protection. When those phrases or sentences are strung together to make a novel or other creative, written work, however, the grouping of emoji may rise to the level of originality necessary for copyright protection. A short string of emoji used to ask a friend to grab pizza and a movie would likely merge with the idea of asking the question in the emoji language, and therefore be uncopyrightable under merger doctrine. Yet, rewriting Moby Dick in emoji requires creative choices and the resulting emoji novel would be copyrightable.[99]

It is possible that short strings of emoji that are unprotectable under copyright could be eligible for trademark protection. Trademark law gives the producer a monopoly on the mark, allowing him or her to prevent competitors from using it.[100] Trademarks can include "any word, name, symbol, or device, or any combination thereof" used by a person "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."[101] Under this definition, it is unlikely that a single emoji or a combination of emoji used commonly in communication could be trademarked because such commonly used emoji would not be able to "identify or distinguish" a brand. This policy allows emoji as words and emoji as sentences or phrases to remain in the public domain as emoji develop as a means of communication. A combination of emoji that would not readily communicate an idea unless associated with a brand, however, may be eligible for trademark protection. For example, a company called 💾 🌵, or Disk Cactus, may be able to register its brand name as a trademark because the combination 💾🌵 is unique and will only have meaning if associated with the brand.[102]

---

[98] *See generally All Emojigrams, A-Z*, EMOJISAURUS, http://emojisaurus.com/all (last visited May 14, 2015).

[99] Allen, *supra* note 11.

[100] Qualitex Co. v. Jacobsen Prod. Co., 514 U.S. 159, 162 (1995) (internal citations omitted).

[101] 15 U.S.C. § 1127.

[102] *See* DISK CACTUS, http://www.diskcactus.com/ (last visited May 14, 2015).

C. Computer Program Protection of Unicode Standard Coding of Emoji

Another category of original, fixed works protected by copyright law is computer programs.[103] The Copyright Act defines a computer program as "a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result."[104] Computer programs qualify as literary works even though they are not written using words.[105] As with PGS works (and all copyrightable works), a computer program must be fixed and original in order to be protected by copyright.[106] Each piece of emoji code is a short combination of letters and numbers.[107] These combinations of letters and numbers tell a computer to pull up the specific emoji the user wishes to type.[108] Therefore, the emoji code technically fits within the copyright definition of a computer program.[109] It is more natural, however, to understand the computer code enabling the use of the emoji, such as the software that allows the user to type with emoji, to be understood as a computer program. Because it is not very natural to think of something that is as short as emoji code as a computer program, it is likely that the code underlying emoji is not copyrightable as a computer program.

The lack of copyright protection for emoji's underlying code would ensure that copyright law could not interfere with the Unicode Consortium's work in making emoji (and all languages) interoperable across digital platforms. Interoperability ensures that people using all types of computer and Internet platforms can communicate with one another via the Internet. Otherwise, it is possible that something typed on an Apple computer or a Google Chrome Internet browser would show up as empty rectangle, rather than readable type, on a Windows computer or an Internet Explorer browser.[110] If every digital platform had to create its own computer code for emoji in order to avoid infringing another platform's copyright, then users on different platforms would never be able to send each other emoji.

---

[103] Apple Comp., Inc. v. Franklin Comp. Corp., 714 F.2d 1240, 1247 (1983); *see* 17 U.S.C. § 102.

[104] 17 U.S.C. § 101.

[105] *Apple Comp.*, 714 F.2d at 1249.

[106] 17 U.S.C. § 102.

[107] *Emoji and Dingbats*, *supra* note 34.

[108] *See id.*

[109] *See* 17 U.S.C. § 101.

[110] *See supra* Part I.A.

D. The Standard-Essential Patent Model and Compulsory Licensing of Emoji under Copyright

It may be tempting to draw parallels between emoji and products that must conform to a standard in order to be usable, such as electronics with outlet-plugs. Emoji require code that conforms to interoperability standards set by the Unicode Consortium, much like a plug must fit a standardized outlet. The plug and outlet are covered by the patent system because they are useful articles. Specific types of plugs and specific types of outlets, however, are required in order for products to function in the U.S., so the patent system cannot give any one entity a monopoly over these standards. Therefore, a special category of patents exists, called standard-essential patents, which require "fair, reasonable, and non-discriminatory" licensing.[111]

Perhaps a similar type of special copyright protection could be given to emoji, requiring compulsory licensing in a similar manner. This would mean that the emoji code created by the Unicode Consortium would be licensed to computer manufacturers for a fair, reasonable, and non-discriminatory rate. The manufacturers who choose to pay the licensing rate could then develop images and keyboards for the emoji of their choice and issue another fair, reasonable, and non-discriminatory license to users who wish to use emoji outside of private communications.[112]

Even with "fair, reasonable, and non-discriminatory" licensing, however, such a system would erect barriers to emoji use that do not seem to be required to fulfill the Framers' intellectual property goals: the promotion of learning and the creation of works that produce "utility."[113] The Unicode Consortium, which controls the creation and management of all emoji code, is a non-profit and would not benefit from a license that would make emoji code into a commercial and profitable good.[114] The companies that create emoji images and make them usable by consumers could potentially make a profit by licensing emoji for public display, but it seems unlikely that this would be a very profitable market. Consumers are

---

[111] Rajendra K. Bera, *Standard Essential Patents (SEPs) and "Fair, Reasonable and Non-Discriminatory" (FRAND) Licensing*, 1 (Jan. 29, 2015) (unpublished article), http://papers.ssrn.com/abstract_id=2557390.

[112] *See* 17 U.S.C. § 106 (Stating private use of a copyrighted work is not prohibited by copyright).

[113] Golan v. Holder, 132 S.Ct. 873, 901 (2012) (Breyer, J., dissenting).

[114] *The Unicode Consortium*, UNICODE, http://www.unicode.org/consortium/consort.html (last visited May 14, 2015).

already accustomed to being able to use emoji however and whenever they wish without the need to pay for a license.[115] A licensing system may not cause users to abandon emoji altogether, especially not in the realm of private communications, but may limit creation of emoji-related resources if the creators of those resources cannot or will not pay licensing fees.[116] Companies could sue users for public display of copyright-protected emoji images, but the companies would be suing their own consumers, which would likely be bad for business and public relations. Moreover, users may be able to bring strong fair use defenses that would result in lengthy and costly litigation.[117] Overall, this system of copyright-with-licensing would create an unnecessarily confusing and complicated marketplace.

An analysis of current U.S. copyright law reveals that it may be possible to protect emoji in some way. Perhaps individual emoji images could be given copyright protection or could be subject to a compulsory licensing scheme. Even if the copyright regime allowed these protections, however, public policy dictates that emoji should be excluded from copyright protection and left to the public domain.

III
EMOJI AS CONSTRUCTED LANGUAGE AND WHY COPYRIGHT CANNOT BE APPLIED

Emoji are more than just a set of small glyphs; they can be seen as a burgeoning means of communication. Part A of this section will introduce the idea of constructed languages and explain how emoji may be seen as a constructed language accessory. Part B will explain how emoji will continue to grow if categorically excluded from copyright protection in order to promote emoji's use as language. Part C will explain how emoji may be governed and developed if left to the public domain.

A. Emoji and Their Development as a Constructed Language Accessory

In many ways, emoji may be best likened to constructed languages. A constructed language, also known as an invented or planned language, is a language owing its origins to an individual human inventor, as opposed to a

---

[115] Hence the creation of resources like EmojiTracker. EMOJITRACKER, *supra* note 10.

[116] Some examples of such resources include the Emoji Dictionary, Emojitracker, and translations of previously existing works into emoji. EMOJI DICTIONARY, http://emojidictionary.emojifoundation.com/home.php?people (last visited May 14, 2015); EMOJITRACKER, *supra* note 10; Allen, *supra* note 11.

[117] *See* 17 U.S.C. § 107. *See, e.g.*, Cariou v. Prince, 714 F.3d 694 (2d. Cir. 2013), Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007).

language that originates and evolves as people communicate over time.[118] While a constructed language may owe its "phonology, morphology, syntax, and sometimes alphabet" to a human inventor, aspects of the language may evolve organically from the inventor's starting point.[119] English, Chinese, and Russian are a few of the many natural languages used today. Examples of constructed languages include Solresol, a language based on music scales developed by Jean François Sudre in the 1830s; Esperanto, a language created by Dr. Ludwig Lazarus Zamenhof in order to help bridge the gap "between hostile groups of Russians, Poles, Germans, and Jews;" and Klingon, a language created for the 1984 film, *Star Trek III: The Search for Spock*.[120]

Like the constructed languages that came before it, emoji were introduced by an individual creator who laid out the bare bones for their use.[121] Since then, of course, emoji have been given life by users who incorporate the symbols into their conversations and writings. This evolution is similar to other constructed languages, such as those created by J.R.R. Tolkien in his *Lord of the Rings* trilogy.[122] After Tolkien laid out the initial vocabulary and structure of Elvish languages such as Quenya and Sindarin, fans studied and expanded the languages in order to write their own works in Tolkien's constructed tongues.[123]

Emoji, however, are not currently developed to the extent of a typical constructed language; as noted above, emoji are not truly a full-fledged language.[124] Emoji have the potential to expand, as Unicode creates additional emoji code and providers illustrate that code, and may eventually constitute a full constructed language.[125] Yet, until then, emoji can be best thought of as a constructed language accessory; a new, invented vocabulary to be added to and mixed with existing language. To some extent, this is what Shigetaka Kurita had in mind when he invented emoji. Kurita wanted Japanese Internet and mobile users to be able to communicate in short-form, without losing the expressiveness seen in

---

[118] Michael Adelman, Constructed Languages and Copyright: A Brief History and Proposal for Divorce, 27 HARV. J.L. & TECH. 543, 545 (2014).

[119] *See id.*

[120] *Id.* at 545, 547-48.

[121] *See generally* Blagdon, *supra* note 1.

[122] Adelman, *supra* note 118, at 551-52.

[123] *Id.*

[124] *See* discussion *supra* Part I.B.

[125] For example, Apple's release of iOS 8.3 on April 8, 2015, added more than 300 new emoji to its selection. Chowdhry, *supra* note 7.

traditional Japanese written letters.[126] What Kurita did not anticipate was that emoji would become popular outside of Japan and find their way into communications worldwide.[127] Therefore, Kurita is the inventor of the words and vocabulary of emoji (the images themselves), but users develop emoji's grammar and fit them into the context of preexisting languages. Emoji were constructed[128] by Kurita,[129] but they are used as an accessory to preexisting natural languages, rather than as a comprehensive language on their own.

The only way for emoji to grow as a meaningful language accessory is for people to use them and figure out how best to incorporate them into their communications. Currently, barriers for using and learning emoji are very low. Generally, all a user has to do is opt to add an emoji keyboard on her smartphone, and she will have all the tools she needs to begin using emoji.[130] A typical constructed language may attract very few people and require time and dedication to learn the language. For example, the most-watched episode of HBO's Game of Thrones had approximately 7.1 million viewers.[131] Of those viewers, only a small percentage will bother to learn the show's constructed languages, Dothraki and Valyrian, especially given their limited vocabularies and functionality.[132] Emoji, however, are available on most smartphones. As of January 2014, approximately 58% of all American adults used a smartphone, in addition to many younger Americans.[133] Any of these smart phone users can follow simple instructions to add an emoji keyboard to her phone's standard set of keyboards and can begin to use emoji without any further knowledge.[134]

Categorically excluding emoji from copyright protection by classifying them as a form of language will keep barriers to using emoji low. The more emoji are used, the more people will come to associate meaning with them and the more emoji will be able to be used for communication. Some users may give up on

---

[126] Blagdon, *supra* note 1.

[127] Negishi, *supra* note 5.

[128] *See* Adelman, *supra* note 118, at 545.

[129] Negishi, *supra* note 5.

[130] *See, e.g.*, Kyli Singh, *How to Enable Emoji on iOS*, MASHABLE (June 17, 2014) http://mashable.com/2014/06/17/emoji-on-ios/.

[131] Rick Kissell, *HBO's 'Game of Thrones' Finale Draws 7.1 Million Viewers Sunday*, VARIETY (June 16, 2014, 8:56 AM), http://variety.com/2014/tv/ratings/hbos-game-of-thrones-closes-with-7-1-million-viewers-sunday-1201221238/.

[132] *See generally* TONGUES OF ICE AND FIRE, http://www.dothraki.org/ (last visited May 14, 2015).

[133] *Mobile Technology Fact Sheet*, PEW RESEARCH CENTER, http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/ (last visited May 14, 2015).

[134] *See, e.g.*, Singh, *supra* note 130.

emoji if they have to purchase materials in order to learn their language or if emoji are not used widely, much like how many fans of television programs featuring constructed languages do not bother to learn those languages. Exclusion of emoji from copyright protection as a form of language will allow them to develop, without hurting the incentives already in place for the creation and maintenance of emoji.

B. Future Growth of Emoji with Categorical Denial of Copyright Protection

It is important to keep barriers to emoji usage low because the more people who use emoji, the more they develop as an accessory to language, giving people a greater range of expression in their communication. As linguist Ben Zimmer stated, "It's the wild west of the emoji era. People are making up the rules as they go. It's completely organic."[135] Therefore, even if the Copyright Office or a court felt that emoji were deserving of copyright protection, society would benefit most if emoji were categorically denied copyright protection.

If emoji are not subject to copyright protection, they will fall into the public domain and be free for all to use. The more people using emoji, the more they may develop as a form of communication. This would promote one of the Founding Fathers' original goals of copyright law: promotion of learning.[136] At the time the Constitution was drafted, the Founder's predominant view was that copyright should be used to encourage the creation of new works in order to improve learning, but should not grant monopolies on intellectual property that would prevent the dissemination of information.[137] In fact, Thomas Jefferson was so fearful that copyright monopolies would stand in the way of learning that he was hesitant to grant copyright and patent protection at all.[138]

Of course, the Founders did decide to include intellectual property protection in the Constitution because they felt, with regard to copyright, that men would to some extent require "encouragement to . . . pursue ideas which may produce utility."[139] It does not appear, however, that men need such encouragement to produce emoji. The first emoji came from Japan, and found their way to the U.S.

---

[135] Alice Robb, *How Using Emoji Makes Us Less Emotional*, NEW REPUBLIC (July 7, 2014), http://www.newrepublic.com/article/118562/emoticons-effect-way-we-communicate-linguists-study-effects.

[136] Golan v. Holder, 132 S.Ct. 873, 901 (2012) (Breyer, J., dissenting).

[137] *Id.*

[138] *Id.*

[139] *Id.*

without copyright protection.[140] Emoji proliferate not only in digital communications, but can be found in the fashion[141] and art worlds.[142] Additionally, despite a lack of copyright protection, new emoji are being created and disseminated. The Unicode Consortium accepts submissions from anyone with an idea for a new emoji.[143] The Consortium evaluates proposals for new emoji and develops suitable proposals into usable characters.[144] Once a Unicode Standard exists for a new character, it is up to the computer platform manufacturers to illustrate the character and make it available on an emoji keyboard.[145] For example, on April 8, 2015, Apple released its iOS 8.3 operating system, which included a new set of emoji that help alleviate the early emojis' lack of diversity, adding more than 300 new emoji to Apple's keyboard.[146]

To exclude emoji from copyright protection under the current U.S. copyright framework, and thereby allow them to develop as a language, emoji would need to be classified as a typeface – i.e. an expression of emoji that itself merges with the idea of the emoji as language. Because typefaces are not copyrightable,[147] typeface classification would mean that users could type with emoji in order to express their ideas and thoughts, and users could display those expressions as desired. Typeface classification would not, however, mean that individual emoji are not copyrightable symbols or pictures.[148] When emoji are not used as language, that is, not used in their typeface capacity but individually as art or adornment,[149] the developer of the emoji display may have a copyright claim based on the individual symbol's copyright protection. Yet, emoji as language would remain in the public domain.

---

[140] Weiss-Meyer, *supra* note 22.

[141] *See, e.g.*, Dennis Green, *Emoji-Inspired Slippers Puts $300 Poo on Your Shoe*, MASHABLE (Mar. 25, 2014), http://mashable.com/2014/03/25/emoji-slippers/, *Emoji-nal Backpack*, NASTY GAL, http://www.nastygal.com/product/emojinal-backpack (last visited May 14, 2015).

[142] *See* EMOJI ART AND DESIGN SHOW, http://www.emojishow.com/ (last visited May 14, 2015).

[143] *Submitting Character Proposals*, UNICODE, (Feb. 12, 2016, 9:53 PM), http://www.unicode.org/pending/proposals.html.

[144] *Id.*

[145] *See* Weiss-Meyer, *supra* note 22.

[146] Chowdhry, *supra* note 7.

[147] Eltra Corp. v. Ringer, 579 F.2d 294, 298 (1978).

[148] Monotype Corp. PLC v. Int'l Typeface Corp., 43 F.3d at n.3.

[149] For example, emoji have been used as adornment on clothing, rather than as a form of communication. *See, e.g.*, CURRENT MOJI, http://www.currentmoji.com/ (last visited May 14, 2015).

C. Governing Emoji as Language without Copyright Protection

Without copyright protection, the greatest number of people can use emoji without creating any disincentive for the continued use of existing emoji or the creation of new characters. This lack of protection raises an important question: who will control the emoji "vocabulary?" Emoji users have developed a unique emoji grammar in a grassroots fashion.[150] Yet, private companies and the Unicode Consortium control the introduction of new characters. Users can only access newly created and existing emoji if computing platform manufacturers choose to make them available on keyboards.[151] The manufacturers, in turn, can only develop images for emoji code provided by the Unicode Consortium.[152] Essentially, the Unicode Consortium and the manufacturers form a de facto language regulator, much like the L'Academie Francaise in France or the Academy of the Hebrew Language in Israel.[153]

It would not make sense for a national governing body, like L'Academie Francaise, to determine the correct use of emoji and when new emoji should be added to the language set. Although emoji technically originated in Japan, they can hardly be considered property of the Japanese government, which had nothing to do with their creation.[154] It would not make sense for the Japanese government to suddenly attempt to reign in and control a language invented by a single citizen and used by people around the world.

The Unicode Consortium seems to have taken the best approach of any player thus far with regard to controlling the development of emoji. Because the goal of Unicode is to make emoji and other characters interoperable, the Unicode Consortium has no reason to favor some digital users over others as it develops new emoji.[155] The Consortium allows submissions from any user who wishes to propose new emoji.[156] The Consortium does not provide much information about

---

[150] Robb, *supra* note 135.

[151] There are currently Unicode Standard codes for 249 national flags, but the flags that show up in a given set of emoji are determined by software manufacturers. For example, Apple's U.S. iPhone features 10 national flags. *Emoji and Dingbats*, *supra* note 34.

[152] *Id.*

[153] *See generally* ACADEMIE FRANCAIS, http://www.academie-francaise.fr/ (last visited May 14, 2015); Dafna Yitzhaki, Minority Languages and Language Policy: The Case of Arabic in Israel, 7 (May 2008) (unpublished Ph.D. thesis, Bar-Ilan University), http://www.academia.edu/11506403/Minority_Languages_and_Language_Policy_The_Case_of_Arabic_in_Israel.

[154] *See* Negishi, *supra* note 5.

[155] *What is Unicode?*, *supra* note 24.

[156] *Submitting Character Proposals*, *supra* note 143.

how emoji proposals are selected for production, other than that technical experts review the proposals.[157] Greater disclosure and transparency of this review process may make people more comfortable with the amount of control the Unicode Consortium has over the development of emoji.

The Unicode Consortium does not decide how emoji will look or which emoji users will be able to access, but perhaps another non-profit organization would be able to fill this gap. By allowing computer manufacturers to fill this gap, the manufacturers are able to shape emoji in a way that may help drive profits, or that may cater to some classes of consumers more than others. In contrast, a disinterested party not driven by profit may be able to read the demands of consumers as a whole and determine the best set of emoji for American users. For example, American users rarely use the Baggage Claim emoji,  ,[158] but would find great use of a dumpling emoji.[159] If a non-profit set an "American emoji keyboard" standard that contained the optimal emoji set, based on research of American consumer demands, computer manufacturers would want to adopt that emoji keyboard for its American products because consumers would prefer the characters in that set. The same logic can be applied to emoji-related software designed for other countries. Furthermore, if the non-profit provided the illustrations for every emoji, even those emoji not provided to an American market could still be operable on American computers because the non-profit will ensure that an image for each emoji code is available worldwide. This emoji-imaging non-profit could work with the Unicode Consortium to illustrate the new emoji code created by the Consortium, thus ensuring that a fair process of emoji development runs from proposal to user-ready character.

CONCLUSION

Emoji were invented to add context and emotion to Japanese consumers' digital communications.[160] Thanks to Apple and other computer platform manufacturers, emoji have become a popular means of communication for U.S. consumers as well.[161] Emoji's rise in popularity in the U.S., however, has come

---

[157] *Id.*

[158] The Baggage Claim emoji is the least-used emoji on Twitter, worldwide. EMOJITRACKER, *supra* note 10.

[159] *See, e.g.*, Ann-Marie Alcantara, *79 New Emoji Are Coming, Including a Dumpling!*, POPSUGAR (Feb. 9, 2016), http://www.popsugar.com/tech/New-Emoji-2016-37539573. The dumpling emoji is a candidate to be coded as an emoji in 2017.

[160] Blagdon, *supra* note 1.

[161] *Id.*

without copyright protections. The American public stands to gain the most from emoji if the characters remain uncopyrightable.

It is possible that individual emoji could be eligible for copyright protection as pictorial, graphic, and sculptural works.[162] When emoji are taken as a group, however, they closely resemble a typeface, and typefaces are categorically excluded from copyright.[163] The code that causes emoji to function on users' devices is also unlikely to be copyrightable because it is too short to truly fit the definition of computer code.[164]

Even if emoji were found to be copyrightable under current U.S. copyright law, public policy dictates that symbols should be categorically excluded from copyright protection. Without copyright protection, emoji will remain in the public domain, and users will be free to use the symbols as a form of expression that will add to the strong American tradition of free speech. To give any entity copyright protection over emoji – or even over one set of emoji illustrations – would create barriers to communication and free expression. As emoji become more ubiquitous in society, users are learning to express themselves through these symbols. Copyright protection of emoji would hamper this growing area of free expression.

People do not need copyright to incentivize the management and creation of new emoji. Both non-profit and for-profit companies are currently working to bring new emoji to users without any promise of intellectual property protections.[165] Excluding emoji from copyright protections favors users, and it is the users who have made emoji into the powerful tool of communication that they have become.

---

[162] *See* 17 U.S.C. § 101, 102.
[163] Eltra Corp. v. Ringer, 579 F.2d 294, 298 (1978).
[164] *See* 17 U.S.C. § 101, 102.
[165] *See Submitting Character Proposals*, *supra* note 141; Chowdhry, *supra* note 7.

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 5 SPRING 2016 NUMBER 2

LET'S BE REASONABLE! THE BROADEST REASONABLE INTERPRETATION IN THE PTAB

JULIAN PYMENTO*

While the Patent Trial and Appeal Board (PTAB), a part of the United States Patent and Trademark Office, uses the "broadest reasonable interpretation" for inter partes review proceedings (IPRs), district courts apply the "ordinary and customary meaning" standard for civil litigation. The disparity between standards in similar adjudicatory proceedings is not explicitly justified by existing law and creates uncertain outcomes for practitioners. This note explores the usage of the broadest reasonable interpretation in IPRs, the problems created by dichotomous claim construction standards between the two forums, and possible avenues of correcting this disparity. The Supreme Court acknowledged the issue posed by disparate standards by granting certiorari in Cuozzo Speed Technologies, LLC v. Lee. Therefore, this note argues that the Supreme Court, in this case, ought to require the PTAB to use the district court's "ordinary and customary meaning" standard.

* J.D. Candidate, New York University School of Law, 2017. The author would like to thank the Editorial Board of the Journal of Intellectual Property & Entertainment Law, particularly Nicole Lieberman, and Notes Program peers.

INTRODUCTION

"The devil is in the details," or so the old adage goes. For American patents, one can easily say that the devil is in the claims, for the claims define the metes and bounds of an invention.[1] Thus, an accurate interpretation of these claims is imperative. The Patent Trial and Appeal Board (PTAB) and district courts currently use disparate claim construction standards, which make claim interpretations ambiguous.

The United States Patent and Trademark Office (USPTO) uses the "broadest reasonable interpretation" standard for all procedures, including patent examination. The PTAB, a part of the USPTO created by the America Invents Act (AIA), also uses the broadest reasonable interpretation.[2] The PTAB applies this standard during post-grant proceedings for patents such as *inter partes* reviews

[1] MPEP § 2173 (9th ed., Nov. 2015) ("It is of utmost importance that patents issue with definite claims that clearly and precisely inform persons skilled in the art of the boundaries of protected subject matter.").

[2] 35 U.S.C. § 6(a) (2011); Dennis Crouch, *BPAI: PTO Should Apply Broadest Reasonable Claim Interpretation to Section 101 Analysis*, PATENTLY-O, (Dec. 17, 2008) http://patentlyo.com/patent/2008/12/bpai-pto-should.html (showing that the Board of Patent Appeals and Interferences, the PTAB's predecessor, also applied the broadest reasonable interpretation).

(IPRs).[3] In contrast to the USPTO and the PTAB, district courts apply the "ordinary and customary meaning"[4] standard when interpreting claims during patent litigation. Both IPRs and civil cases are adjudicatory proceedings, yet there are disparate standards applied across the two. The disparity between the standards creates uncertain outcomes for practitioners and increases the overall amount of litigation in the courts.

In 2015, the Federal Circuit took up the issue, hearing *In re Cuozzo Speed Technologies (*"*In re Cuozzo*"*)* en banc,[5] which challenged the PTAB's standard for evaluating the scope of patent claims. Cuozzo Speed Technologies ("Cuozzo") challenged the PTAB's usage of the broadest reasonable interpretation standard after several of their claims were deemed obvious, arguing instead they should apply the ordinary and customary interpretation standard employed by the district courts.[6] To Cuozzo's dismay, the Federal Circuit upheld the PTAB's standard. Yet, the Federal Circuit was not unanimous. Judge Newman's scathing dissent brought many competing policy concerns to the fore, including the deferential authority granted to the USPTO and the similarities of IPRs to district court proceedings.[7] Adding fuel to the fire, Congress proposed legislation, including the Innovation Act,[8] which seeks, in part, to require the PTAB to use the district court standard. In early 2016, the Supreme Court acknowledged the importance of disparate standards, granting certiorari to Cuozzo's appeal.[9] During oral argument, many of Judge Newman's arguments were brought by Petitioner Cuozzo and weighed by the Court.[10]

Based on the oral argument, the outcome of the Supreme Court decision is still anyone's guess. However, this note argues that the patent system would

---

[3] § 6(b)(4).

[4] 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.").

[5] *In re* Cuozzo Speed Technologies, LLC, 793 F.3d 1268 (Fed. Cir. 2015).

[6] *Id.* at 1275.

[7] *Id.* at 1289 (Newman, J., dissenting) (describing an IPR as "a trial, adjudicatory in nature [which] constitutes litigation") (quoting Google Inc. v. Jongerius Panoramic Techs., LLC, No. IPR2013-00191, at 4 (P.T.A.B. Feb. 13, 2014).

[8] Innovation Act, H.R. 3309, 113th Cong. § 9(b)(1) (2013), https://www.congress.gov/bill/113th-congress/house-bill/3309/text.

[9] Cuozzo Speed Techs., LLC v. Lee, 136 S. Ct. 890 (2016).

[10] *See generally* Oral Argument, Cuozzo Speed Techs., LLC v. Lee, 136 S. Ct. 890 (2016) (No. 15-446), *available at* http://www.supremecourt.gov/oral_arguments/argument_transcripts/15-446_2dp3.pdf.

benefit from aligning the PTAB claim construction standard with the district court standard. In particular, this note concludes that the Supreme Court ought to reverse the Federal Circuit's *In re Cuozzo* decision to alleviate patent owners' fears of defending their patent rights on separate fronts with different claim construction standards.

Part I of this note examines the history of the broadest reasonable interpretation as a claim construction standard within the USPTO. This section explains why the USPTO adopted this interpretation, particularly for patent examination. This section also provides a brief history behind the USPTO's derivation of procedural authority to institute this standard within the PTAB. The section then delves into the how the AIA implemented a new procedure, IPR, and how this procedure quickly rose in popularity. Lastly, this section briefly discusses deference given by the courts and Congress to the USPTO, and how this deference has allowed for the broadest reasonable interpretation as a claim construction standard.

Part II addresses the disparities and similarities between the claim construction standards used in the PTAB and those used in the district courts. This section then expounds upon how these distinct standards impact practice. Specifically, the lack of a uniform standard leads to unnecessary outcome uncertainty. This section also describes how the different standards promote inefficient procedures at both the PTAB and in the district courts. It examines the *In re Cuozzo* case and its holding that the PTAB is entitled to use the broadest reasonable interpretation for post-grant proceedings. Lastly, this section analyzes Judge Newman's dissent and why she believes that the district court standard should prevail even at the PTAB.

Part III first focuses on the the procedural-substantive distinction that has historically affected the USPTO's rulemaking authority and how the Federal Circuit's *In re Cuozzo* decision undermines that distinction. This section weighs congressional intent through the lens of legislation, particularly the Innovation Act, which proposes to remedy post-grant claim construction by replacing the broadest reasonable interpretation with the district court standard. It delves into the USPTO's response to the *In re Cuozzo* decision and the increased difficulty patent owners face to amend claims in IPRs. Conventionally, the ability to liberally amend claims justifies the broadest reasonable interpretation's greater scope. Finally, this section asserts that the Supreme Court ought to reverse *In re Cuozzo* because the USPTO and Congress have not been effecting any meaningful changes since the Federal Circuit case.

I
THE BROADEST REASONABLE INTERPRETATION AND *INTER PARTES* REVIEW

For at least the past century, the USPTO's definitive claim construction standard has been the broadest reasonable interpretation. Alongside post-grant review (PGR) and covered business method (CBM), the *inter partes* review utilizes the broadest reasonable interpretation for claim construction. An IPR is a new procedure created by the AIA to replace the pre-AIA *inter partes* reexamination, providing a means to challenge patent validity that is faster and less costly than civil litigation. Three years after AIA' passage, the PTAB's rate of patent claim survival at the time of final written decision continues to be dismally low.[11] This section argues that a main reason why most patent claims in IPRs are rejected is due to the broadest reasonable interpretation standard.

A. The USPTO's Adoption of the Broadest Reasonable Interpretation

The broadest reasonable interpretation has been the claim construction standard of the USPTO for at least the last century.[12] In addition to taking the broadest interpretation of words in a claim, the USPTO considers their "ordinary usage as they would be understood by one of ordinary skill in the art, taking into account whatever enlightenment by way of definitions or otherwise that may be afforded by the written description contained in the applicant's specification."[13] The Federal Circuit goes on to explain that the broadest reasonable interpretation must be "consistent with the specification" and does not "ignore any interpretative guidance afforded by the applicant's written description."[14]

The foremost justification for the broadest reasonable interpretation is simply that it allows the USPTO to consider the full range of reasonable claim interpretations.[15] Considering this full range is desirable because after patent issuance a district court or patent observer may assume that the Patent Office

---

[11] Matt Cutler, *3 Years of IPR: A Look at the Stats*, LAW 360 (Oct. 9, 2015, 3:59 PM), http://www.law360.com/articles/699867/3-years-of-ipr-a-look-at-the-stats.

[12] *See In re* Carr, 297 F. 542, 543 (D.C. Cir. 1924); *see also In re* Kebrich, 201 F.2d 951, 954 (C.C.P.A. 1953) ("[I]t is very definitely settled by a line of consistent decisions rendered during a long period of time that . . . the tribunals and the reviewing courts in the initial consideration of patentability will give claims the broadest interpretation which, within reason, may be applied."); Phillips v. AWH Corp., 415 F.3d 1303, 1316 (Fed. Cir. 2005) (en banc).

[13] *In re* Morris, 127 F.3d 1048, 1054 (Fed. Cir. 1997).

[14] *Id.*

[15] *In re* Am. Acad. of Sci. Tech Ctr., 367 F.3d 1359, 1364 (Fed. Cir. 2004) (quoting *In re* Yamamoto, 740 F.2d 1569, 1571 (Fed. Cir. 1984)); *see also In re* Prater, 415 F.2d 1393, 1404–05 (C.C.P.A. 1969).

considered any reasonable interpretation of the claims for the purpose of infringement and invalidity defenses.[16] A second and important reason for the USPTO's broadest reasonable interpretation standard is to motivate applicants to narrow their claims through amendments during prosecution. Through claim amendments, the Patent Office recognizes ambiguities, explores scope and breadth of claim language, and imposes clarification.[17] A third justification is the practical one of affording patents a post-issuance presumption of validity in district court proceedings.[18] The presumption of validity makes a district court proceeding more efficient because it does not have to re-determine the upper metes and bounds of the claims. Instead, the broadest reasonable interpretation allows the USPTO's procedures to provide both the public and district courts notice of the scope of the invention before the presumption of validity attaches.[19]

B. Implementation of Inter Partes Review by the AIA

The AIA established a new framework for challenging the patentability of issued patent claims at the USPTO and also created the PTAB for handling those challenges. Overall, the new USPTO proceedings are intended to create a more efficient and streamlined patent system that will improve patent quality and limit unnecessary and counterproductive litigation costs.[20] Presumably, this will promote the climate for investment and industrial activity while simultaneously facilitating the quick removal of subpar patents.[21]

The *inter partes* review is one of the aforementioned PTAB proceedings and it includes a trial that is handled by a panel of three PTAB judges.[22] When IPRs

---

[16] *See* Am. Acad., 367 F.3d at 1364 ("Giving claims their broadest reasonable construction serves the public interest by reducing the possibility that claims, finally allowed, will be given broader scope than is justified.").

[17] *See In re* Zletz, 893 F.2d 319, 321–22 (Fed. Cir. 1989) ("An essential purpose of patent examination is to fashion claims that are precise, clear, correct, and unambiguous. Only in this way can uncertainties of claim scope be removed, as much as possible, during the administrative process.").

[18] *See* cases cited *supra* note 12.

[19] *Id.*

[20] *See* Changes to Implement Transitional Program for Covered Business Method Patents, 77 Fed. Reg. 7080, 7081 (Feb. 10, 2012).

[21] *Id.* at 7092.

[22] 35 U.S.C. §§ 311–319 (2013).

emerged in 2011 as part of the AIA,[23] Congress intended the proceeding to be a "faster, less costly alternative[] to civil litigation to challenge patents."[24]

Prior to 1999, all reexamination proceedings in the USPTO were *ex parte*,[25] meaning that the requestor was precluded from taking part in the proceeding after filing the reexamination request.[26] In 1999, Congress added *inter partes* reexamination, which allowed the requestor to participate in the proceeding through appeal.[27] These *inter partes* reexaminations were often costly and time consuming.[28] The *inter partes* reexamination was similar to the initial patent examination and did not have a settlement option, which would have expedited the process and provided a lower cost.[29] As a result, many courts rejected motions to stay cases pending *inter partes* reexamination even though, from 1999 through June 2012, eighty-nine percent of all requests for reexamination resulted in either total claim cancellation or claim changes.[30] In light of the dilemmas that plagued *inter partes* reexamination, it was no wonder that the AIA sought to replace the procedure with a leaner, faster model.

A post-AIA IPR proceeding is instituted by filing a petition.[31] The petition may challenge the patentability of claims under 35 U.S.C. section 102's novelty requirement or section 103's non-obvious requirement based on prior patents or printed publications.[32] The standard for instituting an IPR proceeding is "a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition."[33] After a petition is filed, the patent owner has three months to file a preliminary response to the petition setting forth reasons why no *inter partes* review should be instituted.[34] The PTAB has three months

---

[23] 157 CONG. REC. S952 (daily ed. Feb. 28, 2011) (statement of Sen. Grassley).

[24] *Id.*

[25] Bayh-Boyle Act, Pub. L. No. 96-517, § 1, 94 Stat. 3015, 3016 (1980) (codified as amended at 35 U.S.C. § 304).

[26] 35 U.S.C. § 305 (2011) ("reexamination will be conducted according to the procedures established for initial examination").

[27] 35 U.S.C. §§ 311, 315; Pub. L. No. 106-113, § 1000(a)(9), 113 Stat. 1501, 1536 (1999).

[28] H.R. REP. NO. 112-98, at 45 (2011), *reprinted in* 2011 U.S.C.C.A.N. 67, 75.

[29] Abbott Labs. v. Cordis Corp., 710 E3d 1318, 1326 (Fed. Cir. 2013) (describing the purpose of switching from *inter partes* reexamination to *inter partes* review "was to convert inter partes reexamination from an examinational to an adjudicative proceeding.").

[30] *See supra* note 28; *see e.g.*, Senorx Inc. v. Hologic, Inc., No. 12-173-LPS-CJB, 2013 WL 144255, at *8–9 (D. Del. Jan. 11, 2013).

[31] 35 U.S.C. § 311(a).

[32] § 311(b).

[33] § 314(a).

[34] § 313.

after the patent owner's response is filed or after the preliminary response is due to decide whether to grant the petition in whole or in part and whether to institute an *inter partes* review proceeding.[35] Once instituted, the PTAB must issue a final, written decision within one year from the date the proceeding is instituted.[36] The decision instituting an IPR will usually include an initial ruling on claim construction. In the response, the patent owner may challenge the petitioner's proposed claim construction and the PTAB's initial claim construction.[37] During *inter partes* review, the patent owner may file one motion to amend the patent as a matter of right, but an amendment may not enlarge the scope of the claims or introduce new matter.[38] The issues decided in an IPR typically have estoppel effect in district court proceedings, thus limiting the invalidity arguments that the patent challenger (or one in privity with the patent challenger) may later invoke in district court.[39] Only the Federal Circuit takes appeals from PTAB final written decisions, making that court the final arbiter of IPRs short of United States Supreme Court review.[40] The AIA's new procedure largely satisfied its objective to make the process faster and more cost effective. In the two years since the *inter partes* review replaced *inter partes* reexamination, 1,800 petitions were filed with the PTAB,[41] and as of August 2014, approximately seventy-five percent of IPR petitions had resulted in the institution of proceedings by the PTAB.[42]

Aside from factors that make IPRs more desirable than the old inter partes reexamination, there are significant differences in the burden of proof in IPR proceedings versus district court cases that may make an IPR more desirable than a civil case for a petitioner. First, unlike in district court cases, in IPR proceedings the patent owner is not entitled to a presumption of validity.[43] Second, in IPR

---

[35] § 314(b).

[36] § 3l6(a)(II).

[37] *See e.g.*, Microstrategy, Inc. v. Zillow, Inc., IPR2013-00034, Paper 42 at 6–7, 20–22 (Mar. 27, 2014) (discussing patent owner's arguments regarding claim construction).

[38] § 316(d)(1), (3).

[39] § 315(e)(2).

[40] § 319.

[41] USPTO, PATENT TRIAL AND APPEAL BOARD: AIA PROGRESS, STATISTICS, 1 (Aug 7, 2014), http://www.uspto.gov/sites/default/files/ip/boards/bpai/stats/aia_statistics_08_07_2014.pdf.

[42] USPTO, AIA PROGRESS, 4 (Aug. 14, 2014), http://www.uspto.gov/ip/boards/bpai/stats/081414_aia_stat_graph.pdf.

[43] § 316(e) (the petitioner shall have the burden of proving a proposition of unpatentability by a preponderance of the evidence); Shaffer, Robert F., Hendrix, Justin A., *Post Grant Proceedings of the AIA Provide New Opportunities and Require Reconsideration of Old Patent Litigation Strategies*, FINNEGAN (June 15, 2012), http://www.finnegan.com/resources/articles/articlesdetail.aspx?news=598696f7-7eba-4fcb-83b8-2369caa91dd3 (petitioner must prove invalidity by a

proceedings the petitioner shall have the burden of proving a proposition of unpatentability by a preponderance of the evidence.[44] In district courts, the accused infringer must prove that the patent is invalid by clear and convincing evidence.[45]

The final difference is the PTAB's use of the broadest reasonable interpretation, under which a pending claim must be "given its broadest reasonable construction in light of the specification."[46] As explained in the published comments to the rules governing *inter partes* review, this standard is "consistent with longstanding established principles of claim construction before the [USPTO]."[47] Where the broadest reasonable construction may differ from that applied by a court is that, through the amendment process, claim coverage can be disavowed throughout the prosecution history. In *Microstrategy v. Zillow*,[48] a patent owner attempted to disavow claim scope in the IPR proceeding in order to narrow the meaning.[49] The PTAB refused, explaining that the patent owner had an opportunity to amend the claim in the same proceeding, and absent such action, the broadest reasonable interpretation should apply.[50]

Obviously, the broadest reasonable interpretation standard is friendly to patent challengers in IPRs because when a claim is interpreted broadly, it is more likely to run into invalidating prior art.[51] It is also no coincidence that IPRs have proven very popular among petitioners. Therefore, IPRs, initiated primarily by petitioners, have seen a steady increase in the years since their inception. IPRs could be requested as early as September 16, 2012; in 2012, a total of ninety-six petitions were filed, in 2013 there were approximately 700, and in 2014 the total

---

preponderance of the evidence, a burden of proof that is much lower than a district court's "clear and convincing" standard where the patent enjoys a presumption of validity.).

[44] § 316(e).

[45] Microsoft Corp. v. i4i Ltd. P'ship, 131 S. Ct. 2238, 2242 (2011) (upholding the clear and convincing evidence standard).

[46] 37 C.F.R. § 42.100(b) (2014).

[47] Changes to Implement Inter Partes Review Proceedings, Post-Grant Review Proceedings, and Transitional Program for Covered Business Method Patents, 77 Fed. Reg. 48,680, 48,688 (Aug. 14, 2012).

[48] No. IPR2013-00034 (P.T.A.B. Mar. 27, 2014).

[49] *Id*. at 10–11.

[50] *Id*. at 12.

[51] Michelle Carniaux & Julia Tanase, *IPR and CBM Statistics*, IPR BLOG (Apr. 7, 2014), http://interpartesreviewblog.com/ipr-cbm-statistics.

more than doubled to about 1,300.[52] These numbers have been far greater than expected and further growth is probable.[53]

C. The USPTO's Rulemaking Authority

Historically, the USPTO's rulemaking authority is embodied in 35 U.S.C. section 2(b)(2).[54] While not explicit, this statute grants only procedural rulemaking authority.[55] *Chevron* deference, as per *Chevron, U.S.A., Inc. v. Natural Resources Defense Council, Inc.*,[56] applies only where Congress confers rulemaking authority.[57]

The AIA, enacted in 2011, arguably provides additional grants of rulemaking authority outside of section 2(b)(2) because it gives the USPTO authority to promulgate rules that implement new trial-like proceedings, such as IPRs.[58] Specifically, for IPRs, the USPTO's rulemaking duties include "setting forth the standards for the showing of sufficient grounds to institute a review"[59] and "establishing and governing" the review.[60] Additionally, the AIA gives the USPTO the authority to set its own fees,[61] and permits the agency to make certain policy

---

[52] USPTO, PATENT TRIAL AND APPEAL BOARD STATISTICS, 4 (Apr. 30, 2015), http://www.uspto.gov/sites/default/files/documents/2015-04-30%20PTAB.pdf.

[53] Changes to Implement Inter Partes Review Proceedings, 77 Fed. Reg. at 48724 (Initial Patent Office reports estimated approximately 460 petitions would be requested per year. Through only the first four months of fiscal year 2015, 556 petitions for IPRs were filed.).

[54] 35 U.S.C. § 2(b)(2)(A)–(D) (2012) ("The Office . . . may establish regulations, not inconsistent with law, which . . . shall govern the conduct of proceedings in the Office[,] . . . shall facilitate and expedite the processing of patent applications, [and]. . . may govern the recognition and conduct of agents, attorneys, or other persons representing applicants or other parties before the Office . . . .").

[55] Animal Legal Defense Fund v. Quigg, 932 F.2d 920, 930 (Fed. Cir. 1991) ("A substantive declaration with regard to the Commissioner's interpretation of the patent statutes . . . does not fall within the usual interpretation . . . ." of what is now § 2(b)(2)).

[56] 467 U.S. 837 (1984).

[57] United States v. Mead Corp., 533 U.S. 218, 226–27 (2001) ("[A]dministrative implementation of a particular statutory provision qualifies for *Chevron* deference when it appears that Congress delegated authority to the agency generally to make rules carrying the force of law, and that the agency interpretation claiming deference was promulgated in the exercise of that authority."); *Id.* at 842–43 (When a court reviews an agency's rulemaking authority, it must first determine if Congress has spoken directly to the question at issue and second, if there is statutory ambiguity, if the agency's conduct is permissible.).

[58] America Invents Act, § 6(a) (codified at 35 U.S.C. §§ 311–319) (directing USPTO to issue regulations implementing inter partes review).

[59] § 316(a)(2).

[60] § 316(a)(4).

[61] AIA § 10.

judgments about prioritizing patent applications.[62] Therefore, some thought that the institution of *inter partes* review broadened the USPTO's rulemaking authority.[63] Others concluded that Congress would not broaden the USPTO's authority without express language.[64] Whether utilization of the broadest reasonable interpretation falls under such a procedural or substantive distinction is still ambiguous. However, thus far, the broadest reasonable interpretation is definitively what the USPTO uses in all proceedings for unexpired patents.

II
A TALE OF TWO CONSTRUCTION STANDARDS: PTAB & DISTRICT COURTS

The PTAB's standard of claim construction for *inter partes* review creates a dilemma because it directly contrasts with the district courts' standard. While the PTAB uses the broadest reasonable interpretation, district courts utilize the "ordinary and customary meaning" standard laid out in *Phillips v. AWH Corp.*[65] This section highlights the disparity of claim construction standards and discusses the general shortcomings that plague the broadest reasonable interpretation as a whole. Additionally, this section highlights that a dichotomy of standards can lead to varying validity outcomes across arenas of adjudication. The fact that the same issue can result in alternative judgments creates inefficiency in the court system, and greater risk for patentees.

A. Ordinary and Customary Meaning: The District Court Standard

In contrast to the broadest reasonable interpretation standard at the USPTO, district courts construe patent claims according to the *Phillips* standard, which requires claims be given "the meaning that [a] term would have to a person of ordinary skill in the art in question at the time of the invention."[66] This determination is based on the entire record before the court, taking into consideration both intrinsic evidence such as the claims, specification and prosecution history, and extrinsic evidence such as dictionary definitions and expert testimony.[67] The Federal Circuit has made clear that the broadest reasonable

---

[62] AIA § 25 (codified at 35 U.S.C. § 2(b)(2)(G)).

[63] *See generally* Sarah Tran, *Patent Powers*, 25 HARV. J.L. & TECH. 609, 631-35 (2011-2012) ("The creation of post-grant review provides the USPTO with a key opportunity to set substantive patent law standards and make patent policy.").

[64] John M. Golden, *The USPTO's Soft Power: Who Needs Chevron Deference?*, 66 SMU L. REV. 541, 545 (2013) ("I am skeptical that the [AIA] has worked such a sea change through implicit, rather than express, provision.").

[65] Phillips v. AWH Corp., 415 F.3d 1303, 1313 (Fed. Cir. 2005) (en banc).

[66] *Id.* at 1313.

[67] *Id.* at 1315–17.

interpretation must be applied by the USPTO in office proceedings and that the *Phillips* standard must be used by district courts in validity and infringement actions.[68] This dichotomy is so entrenched that the USPTO will use the broadest reasonable interpretation even for the claims that a district court previously construed, and is not required to adopt the district court's construction.[69]

One example where the PTAB's proposed claim constructions of terms differed from the previous definition provided by the district court is *Rackspace Hosting v. Rotatable Tech.*[70] The PTAB and the Eastern District of Texas both determined that "computer display window" was a disputed claim limitation in one of the patent claims.[71] The district court defined the term "computer display window" to mean "a graphic user interface (GUI) displayable on a monitor or a screen" based on the specification.[72] The PTAB noted that the specification did not provide an express definition of the term, but rather relied on the understanding of the term "window" at the time of the invention as shown in a technical dictionary, and concluded that the proper definition was "a division of a display screen in which a set of information is displayed."[73] The result of differing interpretations is that certain prior patents or publications may serve to invalidate one definition but not the other depending on the venue of adjudication. Because the PTAB's claim construction standard is broader, the petitioner is afforded a greater advantage in a post-grant proceeding resulting from the greater breadth of applicable prior art. The PTAB also provides other extraneous petitioner benefits, such as quicker resolution and avoidance of steep litigation costs.

B. The Enigmatic Broadest Reasonable Interpretation

Congress created the Federal Circuit in 1982 to establish uniformity in the substantive patent law applied in USPTO proceedings and district court patent

---

[68] MPEP § 2111 ("Patented claims are not given the broadest reasonable interpretation during court proceedings involving infringement and validity, and can be interpreted based on a fully developed prosecution record.")

[69] *In re* NTP, Inc., 654 F.3d 1268, 1274 (Fed. Cir. 2011); *see also* Jack Henry v. Datatreasury Corp., No. CBM2014-00056, at 6 (P.T.A.B. Jul. 10, 2014) ("Petitioner argues that the district court's interpretation should be adopted, but provides no persuasive analysis as to how the term is to be interpreted under the broadest reasonable interpretation standard, which is different from the standard used by a district court.").

[70] No. IPR2013-00248 (P.T.A.B. Oct. 1, 2013).

[71] *Id.* at 4–5; Rotatable Tech. LLC v. Nokia, 2013 WL 3992930 *6 (E.D. Tex. Aug. 2, 2013).

[72] *Rotatable Tech.*, 2013 WL 3992930, at *6.

[73] *Rackspace*, No. IPR2013-00248, at pp. 4–5.

litigation proceedings.[74] Given that claim construction is often dispositive in infringement actions, uniformity in overall infringement outcomes requires the Federal Circuit to maintain predictable claim interpretations.[75] Prior to 1982, different standards could apply depending on the circuit because each of the appellate courts independently reviewed decisions from their corresponding district courts.[76] The broadest reasonable interpretation standard disrupts the horizontal equity that the creation of the Federal Circuit aimed to establish. By introducing an interpretation methodology unique to the USPTO, the Federal Circuit applies claim interpretation doctrine in USPTO appeals that contradicts the claim interpretation used in district court appeals.[77] This variation in interpretation methodology makes it possible for patent claims to have variable meanings in the Federal Circuit depending on the venue of the original claim.

Not only does the broadest reasonable interpretation undermine patent law uniformity, it increases the risk of incorrect rejections. The same claim that is valid under district court interpretation methodology could be invalid under the broadest reasonable interpretation, which has a strong possibility of Type II errors, false negatives, or in this context, falsely triggering an unpatentability finding.[78] Yet, the system has extensive mechanisms[79] for correcting Type I errors, false positives, which in this context is a false finding of patentability.[80] The USPTO and the courts can remedy Type I errors through reexamination, reissue, or invalidation.[81]

The third and perhaps most obvious criticism of the broadest reasonable interpretation standard is that it is ambiguous. The Manual of Patent Examining

---

[74] *See* Commission on Revision of the Federal Court Appellate System, *Structure and Internal Procedures: Recommendations for Change*, 67 F.R.D. 195, 369–71, 373–74 (1976).

[75] *See, e.g.*, Giles S. Rich, *The Extent of the Protection and Interpretation of Claims--American Perspectives*, 21 INT'L REV. INDUS. PROP. & COPYRIGHT L. 497, 499 (1990) ("[T]he name of the game is the claim").

[76] *See* Rochelle Cooper Dreyfuss, *The Federal Circuit: A Case Study in Specialized Courts*, 64 N.Y.U. L. REV. 1, 6–8 (1989).

[77] *See* Joan E. Schaffner, *Federal Circuit "Choice of Law": Erie Through the Looking Glass*, 81 IOWA L. REV. 1173, 1196 (1996).

[78] Dawn-Marie Bey & Christopher A. Cotropia, *The Unreasonableness of the Patent Office's "Broadest Reasonable Interpretation" Standard*, AIPLA QUARTERLY JOURNAL, Volume 37, Number 3, 285, 304 (Summer 2009).

[79] *See* Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495, 1500 (2001) (pointing out that litigation may be a better way to weed out bad patents).

[80] *Supra* note 78, 291–292.

[81] 35 U.S.C. § 251 (2006) (allowing for the reissue of defectively issued patents); *id.* § 282 (providing district courts with the ability to invalidate claims); *id.* § 301 (setting forth the basis for *ex parte* reexamination proceedings).

Procedure (MPEP) does not provide much guidance as to the bounds of the broadest reasonable interpretation. The MPEP also does not provide examiners with examples, tests, or directions on how to implement this standard.[82] The case law cited in the MPEP exemplifies general claim interpretation principles that apply during both prosecution and enforcement proceedings, but it does not elaborate on the USPTO's unique application.[83]

C. Staying Co-Pending Litigation & Forum Shopping

There are limited instances where both the PTAB and a district court have construed a claim, largely because post-AIA district courts are willing to stay co-pending patent litigation while post-grant proceedings are underway.[84] To the extent courts stay concurrent litigation in favor of a pending inter partes review, the savings for both parties is substantial, and by virtue of the fact that there is one less case on the district court docket, it saves the court time.[85] In 2015, the USPTO estimated the average cost of patent litigation to be $3.1 million where the damages fell between $10 million and $25 million and $600,000 where the damages were less than $1 million.[86] In contrast, an *inter partes* trial proceeding all the way through appeal is estimated to cost about $350,000.[87] The reduced costs can be attributed partly to reduced discovery,[88] and the accelerated timeline Congress imposed on the PTAB to resolve these proceedings.[89]

The legal standard for motions to stay varies slightly between jurisdictions but there are generally three factors that are considered: "(1) whether a stay would unduly prejudice or present a clear tactical disadvantage to the non-moving party;

---

[82] *See* MPEP § 2111.

[83] *Id.*

[84] Robert Arcamona & David Cavanaugh, *Stays to Litigation Pending IPR and CBM Review: Statistics, Trends, and Key Issues*, INTELL. PROP. TODAY 9 (Mar. 2014), https://www.wilmerhale.com/uploadedFiles/Shared_Content/Editorial/Publications/Documents/IP- today-stays-to-litigation-pending.pdf (noting that as of 2014, statistics suggest a rate of stay running at about sixty percent).

[85] 2015 Report of the Economic Survey, AMERICAN INTELLECTUAL PROPERTY LAW ASSOCIATION, 37–38 (June 2015), http://files.ctctcdn.com/e79ee274201/b6ced6c3-d1ee-4ee7-9873-352dbe08d8fd.pdf.

[86] *Id.* at 37.

[87] *Id.* at 38.

[88] *See* 35 U.S.C. § 316(a)(5) (2013) ("[D]iscovery shall be limited to– (A) the deposition of witnesses submitting affidavits or declarations; and (B) what is otherwise necessary in the interest of justice . . . .").

[89] *See* § 316(a)(11) (requiring that the final determination in an inter partes review be issued within 1 year of institution or within 18 months for good cause).

(2) whether a stay will simplify the issues in question and trial of the case; and (3) whether discovery is complete and whether a trial date has been set."[90] Despite the commonality of these factors in many forums for patent litigation, grant rates vary significantly from one forum to another. In the District of Delaware, for example, Judge Sleet has granted nine out of ten motions to stay.[91] In contrast, the Eastern District of Texas has granted fewer than fifty-five percent of the requests.[92] Strikingly, courts in that district grant fewer than thirty percent of requests when the parties dispute the motion to stay.[93]

While the Federal Circuit ordered the Eastern District of Texas to stay litigation in favor of a covered business method review proceeding,[94] this has not occurred for the *inter partes* review. Therefore, there is a greater likelihood of forum shopping in locations like the Eastern District of Texas, which is already notorious for attracting patent owners, specifically patent trolls.[95] In 2002, the Eastern District of Texas had thirty-two patent cases filed.[96] By 2013, the number skyrocketed to 1,495, constituting nearly a quarter of all patent cases filed in the U.S.[97] With a unification of PTAB and district court standards for claim construction, district courts such as the Eastern District of Texas may be more compelled to grant requests to stay litigation, thereby improving the efficiency of the courts' relationship with the PTAB.

D. The Difficulty of Amendments in IPRs

Part I demonstrates that a major part of the broadest reasonable interpretation's rationale is the patentee's ability to amend within the USPTO in order to overcome overly broad interpretations. 35 U.S.C. § 316(d), which has been in place since September 16, 2011, grants patent owners the ability to file a

---

[90] Xerox Corp. v. 3Com Corp., 69 F. Supp. 2d 404, 406 (W.D.N.Y. 1999).

[91] Aashish Kapadia, *Inter Partes Review: A New Paradigm in Patent Litigation*, 23 TEX. INTELL. PROP. L.J. 113, 131–32 (2015).

[92] *Id.*

[93] *Id.*

[94] *See* VirtualAgility Inc. v. Salesforce.com, Inc., 759 F.3d 1307, 1320 (Fed. Cir. 2014) (holding that a district court abused its discretion when it denied a stay pending covered business method review).

[95] Daniel Nazer & Vera Ranieri, *Why Do Patent Trolls Go to Texas? It's Not for the BBQ*, ELECTRONIC FRONTIER FOUNDATION (Jul. 9, 2014), https://www.eff.org/deeplinks/2014/07/why-do-patent-trolls-go-texas-its-not-bbq.

[96] Julie Creswell, *So Small a Town, So Many Patent Suits*, N.Y. TIMES, Sept. 24, 2006.

[97] Owen Byrd & Brian Howard, *Patent 2013 Litigation Year in Review*, Lex Machina, i (May 13, 2014), https://www.law.berkeley.edu/files/2013_Patent_Litigation_Year_in_Review_Full_Report_(MLex_Machina).pdf.

motion to amend their claims during an inter partes review.[98] By right, parties can make one motion to amend and motions can be made later for cause. However, as of June 15, 2015, the PTAB allowed motions to amend in only four IPR proceedings[99] out of the 3,400 IPR petitions that have been filed since the AIA took effect in September 2012.[100] In practice, the right to amend is virtually non-existent.

In *International Flavors & Fragrances Inc. v. United States*,[101] the patent owner finally persuaded the PTAB to grant in part a motion to amend, which illustrates the difficulty of amending within IPRs.[102] In examining the patentability of the proposed substitute claims, the PTAB reiterated that the patent owner bears the burden of proof to demonstrate patentability of the proposed claims over the prior art in general.[103] Even though the petitioner, International Flavors, did not oppose the motion to amend, the PTAB found that the U.S. met its burden of proof for some of the claims but still only granted the United States' motion to amend *in part*.[104]

One may argue that using a standard other than broadest reasonable interpretation in IPRs would create inconsistencies with other USPTO proceedings such as patent examination. Take, for instance, if the district court standard were applied to a patent in an IPR proceeding and then that same patent or another patent in its family goes back to prosecution or a reissue proceeding. This creates an uncertainty and risk for patent owners. However, this line of reasoning is undermined since amendments are freely permitted during prosecution, while it is practically impossible to amend during IPR.

E. In re Cuozzo: Judge Newman's Parry of the Majority Holding

The PTAB's broadest reasonable interpretation standard was one of the foremost issues in *In re Cuozzo Speed Techs.* Cuozzo was the assignee of U.S. Patent No. 6,778,074 ("'074 patent"), entitled "Speed Limit Indicator and Method

---

[98] 35 U.S.C. § 316(d) Amendment of the patent. (Added Sept. 16, 2011, P.L. 112-29, §6(a), 126 Stat. 302).

[99] *See PTAB Allows Motion to Amend in IPR Challenging Neste Oil Patent*, MANAGING INTELLECTUAL PROPERTY, http://www.finnegan.com/news/newsdetail.aspx?news=d0dl8aee-10e8-4511-bfa5-3b2962dala39 (June 15, 2015).

[100] U.S. Patent and Trademark Office, Patent Trial and Appeal Board Statistics 8/3122015, http://www.uspto.gov/sites/default/files/documents/2015-08-31%20PTAB.pdf.

[101] IPR2013-00124, at 7 (PTAB May 20, 2014 (Paper 12).

[102] *Id.* at 7.

[103] *Id.* at 9–10.

[104] *Id.* at 10–11.

for Displaying Speed and the Relevant Speed Limit," which issued on August 17, 2004.[105] The '074 patent's claim 10 is for a "speed limit indicator comprising…a speedometer *integrally attached* to said colored display."[106] The PTAB's broadest reasonable interpretation construed the term "integrally attached" as meaning "discrete parts physically joined together as a unit without each part losing its own separate identity."[107] Under this interpretation, the Board found this claim and two other dependent claims, 14 and 17, unpatentable over the prior art.[108] The Board also denied Cuozzo's motion to amend the patent because the substitute claims lacked written description support."[109] The Board also held that the "substitute claims would improperly enlarge the scope of the claims as construed by the Board."[110]

Cuozzo appealed to the Federal Circuit. The Federal Circuit affirmed the Board's final determination, "finding no error in the Board's claim construction under the broadest reasonable interpretation standard, the Board's obviousness determination, and the Board's denial of Cuozzo's motion to amend."[111] Ultimately, the Federal Circuit found the USPTO has the appropriate authority to use the broadest reasonable interpretation standard in post-grant proceedings.

1. Broadest Reasonable Interpretation: The Law of Claim Construction?

The majority noted that the USPTO applied the broadest reasonable interpretation standard for more than a century in various types of USPTO proceedings from "initial examinations, interferences, and post-grant proceedings such as reissues and reexaminations."[112] Thus, the majority believed that Congress did not design the AIA to change the broadest reasonable interpretation since Congress legislated knowing of its prevailing use in the USPTO, therefore, implicitly approving the existing rule.[113]

Judge Newman, in her dissent, agreed that the broadest reasonable interpretation is authorized for use in the examination of pending applications. However, Newman noted that the standard exists only "to restrict or clarify the

---

[105] U.S. Patent No. 6,778,074 (issued Aug. 17, 2004).
[106] *Id.* col. 7 ll. 1-10.
[107] *In re* Cuozzo Speed Techs., LLC, 793 F.3d 1268, 1272 (Fed. Cir. 2015).
[108] *Id.*
[109] *Id.*
[110] *Id.*
[111] *Id.* at 1271.
[112] *Id.* at 1276.
[113] *Id.* at 1277.

applicant's proposed claims, not to broaden them."[114] Newman argued that the broadest reasonable interpretation ought to be used only in examination, to give the public notice of a patent claim's upper bound.[115] Therefore, the PTAB contradicts the purpose of the broadest reasonable interpretation by extending the broadest reasonable interpretation to post-grant validity, which has generally been held to be the province of the district court.[116]

2. Restrictive Permission to Amend Claims

Cuozzo argued that earlier judicial decisions that utilized the broadest reasonable interpretation relied on the easy availability of seeking claim amendments, but the AIA limits amendments in IPR proceedings. The majority asserted that there are no restrictions on "amendment opportunities that materially distinguish IPR proceedings from their predecessors in the patent statute."[117] For support, the majority cited section 316(d)(1) which provides that a patentee may file one motion to amend in order to "[c]ancel any challenged patent claim" or "[f]or each challenged claim, propose a reasonable number of substitute claims."[118] Therefore, the opportunity to amend, however cabined, justifies using the broadest reasonable interpretation.[119]

The dissent pointed out that the opportunity to amend is completely denied in an infringement action in district court. Even in reexamination proceedings for expired patents, the USPTO applies the district court claim construction standard rather than the broadest reasonable interpretation, because claims cannot be amended in expired patents.[120] Thus, the dissent argued that amendments are provided for the sole purpose of promoting "a fluid exchange between the examiner and the applicant" during examination-type proceedings.[121] The PTAB should not apply the broadest reasonable interpretation when it has effectively removed the applicant's right to amend.[122]

---

[114] *Id.* at 1286 (Newman, J., dissenting).
[115] *Id.* at 1286–87.
[116] *Id.* at 1287.
[117] *Id.* at 1277.
[118] *Id.*
[119] *Id.* at 1278.
[120] *Id.* at 1287 (Newman, J., dissenting).
[121] *Id.*
[122] *Id.* at 1287–88.

3. The IPR Is an Adjudicatory Process

The majority in *In re Cuozzo Speeding Technologies* did not believe that distinguishing an IPR as an adjudicatory process rather than as an examination makes any difference in congressional approval for the broadest reasonable interpretation. After all, interference proceedings are also in some sense adjudicatory, but yet use the broadest reasonable interpretation.[123]

Newman's dissent perceived the adjudicatory versus examination dichotomy to lie at the crux of replacing the old *inter partes* reexamination with the *inter partes* review. Newman notes that the benefits of adversary participation in IPRs, such as providing for discovery, witnesses, argument, and other litigation procedures, are specifically meant to achieve the very same benefits of validity proceedings in the district courts.[124] Newman proceeded to quote the PTAB's own words, saying "'[a]n *inter partes* review is neither a patent examination nor a patent reexamination,' but is 'a trial, adjudicatory in nature [which] constitutes litigation.'"[125] Therefore, Newman found that the USPTO's usage of the pre-grant examination claim construction standard is "curious" and "a negation of the purpose and obligation of this new adjudicatory process."[126] She continued to warn that if the IPRs are not meant to provide a surrogate forum for district courts, then the "new procedures will become no more than a tactical vehicle for delay, harassment, and expenditure," thus fulfilling a congressional warning that the AIA's changes "are not to be used as tools for harassment or a means to prevent market entry through repeated litigation and administrative attacks on the validity of a patent."[127]

4. Agency Rulemaking Authority

The majority also noted that the provisions 35 U.S.C. section 316(a)(2) and (a)(4), mentioned in Part I of this note,[128] convey rulemaking authority to the USPTO to "prescribe regulations," "setting forth the standards for the showing of sufficient grounds to institute . . . review … establishing and governing inter partes

---

[123] *Id.* at 1278.

[124] *Id.* at 1288–89 (Newman, J. dissenting).

[125] *Id.* at 1289 (quoting *Google Inc. v. Jongerius Panoramic Techs., LLC*, No. IPR2013–00191, at 4 (P.T.A.B. Feb. 13, 2014)).

[126] *Id.* at 1289.

[127] *Id.* at 1289 (quoting H.R. REP. NO. 112–98, pt. 1, at 48 (2011), 2011 U.S.C.C.A.N. 67, 78).

[128] *See supra* Part I.C.

review . . . and the relationship of such review to other proceedings . . . ."[129] Because of this Congressional authorization, the majority stated that the *Chevron* framework should apply. This involves determining "whether Congress has directly spoken to the precise question at issue" and, if not, then determining "whether the agency's interpretation is based on a permissible construction of the statutory language at issue."[130] Because Congress was silent on the subject of how the USPTO should construe the "claim," step one of *Chevron* is satisfied and the broadest reasonable interpretation is a "permissible construction of the statutory language," and step two is unnecessary.[131] The broadest reasonable interpretation is a permissible construction because the USPTO merely embodied the approach it has uniformly applied in interpreting claims for pre-IPR examination proceedings.[132] The majority, however, stated that "[w]e do not draw that conclusion from any finding that Congress has newly granted the [USPTO] power to interpret *substantive* statutory 'patentability' standards."[133]

The dissent maintained that there is a strict distinction between procedural and substantive rulemaking here. Newman stated that the "*conduct* of inter partes review" noted in 35 U.S.C. section 316(a)(4) specifically connotes it is a procedural power "including public access to proceedings, discovery rules, and the right to a hearing."[134] Claim construction is a matter of law and therefore, is a substantive matter.[135] As such, claim construction rules do not lie under the procedural rulemaking covered by the statute. Moreover, Newman insisted that "deference is constrained by the obligation to honor the clear meaning of a statute, as revealed by its language, purpose, and history."[136] Taking the obvious purpose of the AIA to be the creation of a surrogate for district court litigation, the USPTO's promulgation of 37 C.F.R. section 42.300(b) authorizing and requiring the broadest reasonable interpretation for *inter partes* reviews is a direct contradiction.[137]

The Federal Circuit's *In re Cuozzo* decision caused consternation for many patent owners who could relate to Cuozzo. However, the en banc decision was split 6-5 and Judge Newman's dissent tactfully countered the majority's arguments,

---

[129] *Id.* at 1275.
[130] *Id.* at 1279.
[131] *Id.*
[132] *Id.*
[133] *Id.*
[134] *Id.* at 1290 (Newman, J., dissenting) (emphasis added).
[135] *Id.* at 1286.
[136] *Id.* at 1290.
[137] *Id.*

lending much credence to a shift in the PTAB's claim construction standard. As such, Cuozzo was able to petition the Supreme Court to obtain certiorari and during oral argument, many of Newman's arguments could be echoed particular in proclaiming IPRs as a surrogate for district court litigation.[138]

III
POST-CUOZZO AND THE FUTURE OF THE BROADEST REASONABLE INTERPRETATION

While the Federal Circuit's *In re Cuozzo* decision ultimately held that the PTAB is entitled to use the broadest reasonable interpretation as the standard for post-grant proceedings, there are many efforts currently underway for change in line with Judge Newman's dissent. First, given the historic lack of deference given to substantive rules by the USPTO, it is unclear that courts should grant the USPTO *Chevron* deference in the matter of the PTAB's post-grant proceeding claim construction standard. Second, the USPTO, through proposed rules, is unlikely to affect any change to the broadest reasonable interpretation in the PTAB or to the underlying problem of virtual inability to amend claims. Third, Congress's silence, knowing that the broadest reasonable interpretation is prevalent in the USPTO, does not necessarily indicate acceptance with regard to its application in the PTAB. This section argues that pending legislation, such as the Innovation Act, are indicative of a concerted effort to adopt Judge Newman's proposal of uniform claim construction approach in both district courts and the PTAB. Finally, the Supreme Court heard oral argument on *In re Cuozzo* on April 25, 2016 and offers the best opportunity for rectifying the claim construction argument once and for all as opposed to the incessant back and forth for which Congress is notorious in passing key legislation.

A. The USPTO's "New" Rulemaking Authority

The disagreement between the majority and dissent in *In re Cuozzo Speeding Technologies* casts a shadow of doubt on whether the AIA granted the USPTO greater rulemaking authority. If so, is the decision to implement the broadest reasonable interpretation standard in the PTAB a substantive or procedural rule? Despite the obvious question, the USPTO promulgated the rule without characterizing it. However, in *SAP America, Inc. v. Versata Development Group, Inc.*,[139] the USPTO justified the broadest reasonable interpretation standard by

[138] Oral Argument at 4:5-8, Cuozzo Speed Techs., LLC v. Lee, 136 S. Ct. 890 (2016) (No. 15-446), *available at* http://www.supremecourt.gov/oral_arguments/argument_transcripts/15-446_2dp3.pdf.

[139] 107 U.S.P.Q.2d 1097, 2013 WL 3167735 (Jun. 11, 2013).

asserting that the AIA "provides the Office with authority exceeding that of merely setting forth 'procedures,'" suggesting it sees its authority as containing substantive rulemaking to some degree.[140] The Federal Circuit said that a USPTO rule is substantive when it "effects a change in *existing law* or policy which affects individual rights and obligations."[141] The Supreme Court stated clearly that claim construction is a matter *of law*, which obviously is outcome determinative.[142]

There are three routes through which this issue of authority can be rectified. The first is through a clear and conclusive declaration by Congress to give the USPTO obvious substantive rulemaking authority, thereby rejecting the long-standing procedural-substantive distinction that historically applied to the USPTO. Another route is if the USPTO pushes its limits on the procedural-substantive dichotomy in further cases at the Federal Circuit level. Finally, the Federal Circuit could either explicitly hold that the broadest reasonable interpretation is a purely procedural issue, or as scholar Allyson Mackavage has recommended, "recognize that the USPTO had at least limited substantive rulemaking authority under the new provisions added by the AIA."[143] Considering the favorable result the USPTO received from the Federal Circuit in *In re Cuozzo*, it is conceivable that the Court might slowly but steadily broaden the definition of procedural rulemaking authority or achieve increased recognition of substantive rulemaking authority. In both scenarios, it is up to the Federal Circuit to make a concerted effort to specify the scope of the USPTO's rulemaking authority regarding the procedural-substantive distinction. Neither option is advantageous over another but the end result ought to be accomplished: clearly delineating the USPTO's rulemaking authority post-AIA.

B. The Legislative Response

Another mechanism through which the PTAB's claim construction standard may be changed is through Congressional action. For example, the Innovation Act of 2015 (H.R.9), introduced by Rep. Bob Goodlatte, if passed, "would override the patent office and explicitly mandate that the *Phillips* standard be used in *inter*

---

[140] *Id.* at 1104.

[141] Cooper Techs. Co. v. Dudas, 536 F.3d 1330, 1336 (Fed. Cir. 2008) (internal quotation marks omitted).

[142] Teva Pharm. USA, Inc. v. Sandoz, Inc., 135 S. Ct. 831, 837 (2015) ("[I]t was proper to treat the ultimate question of the proper construction of the patent as a question of law").

[143] Allyson E. Mackavage, *One-Off or a Sign of Things to Come? In re Cuozzo and the Scope of the United States Patent and Trademark Office's Rulemaking Authority*, 115 COLUM. L. REV. SIDEBAR 93, 102–03 (2015).

partes review."[144] Other bills in the Senate are the PATENT Act (S.1137), introduced by Sen. Chuck Grassley, and the STRONG Patent Act (S.632), introduced by Rep. Christopher Coons, both of which accomplish the same task.[145]

Legislation can be a challenging avenue for accomplishing meaningful change in the claim construction standard because of the compromising nature of the legislative process. The PATENT Act and Innovation Act are good examples of the occasionally conflicting aims of patent legislation. Both the PATENT Act and Innovation Act primarily involve changes to patent litigation, aimed at curbing abusive suits by patent trolls.[146] Patent trolls first obtain patent rights and then threaten to enforce those rights on suspected infringers.[147] In order for the other party to avoid patent litigation, the patent trolls seek settlements and licensing fees to practice the patented invention.[148] Therefore, it is ironic that Congress added the broadest reasonable interpretation provision in proposed legislation at the behest of patent owners, which may include trolls, who were concerned about how frequently the PTAB invalidates patents. These provisions are outliers in an otherwise non-patentee-friendly bill, "since it would tend to make patents owned by non-practicing entities more likely to survive AIA reviews."[149]

Despite the conflicting aims of the Innovation Act, its provision on PTAB claim interpretation align with the views of Judge Newman and patent provisional. The Innovation Act, passed the House of Representatives in 2014 but stalled in the Senate.[150] The bill was then reintroduced early 2015,[151] and specifically requires the

---

[144] David Soofian & Victoria Reines, *Is Broadest Reasonable Interpretation Here To Stay?*, LAW 360 (Apr. 11, 2016, 10:36 PM), http://www.law360.com/articles/677020.

[145] *Id.*; BAKER BOTTS, *Proposed Patent Reform in 2015: The Patent Act & The Innovation Act* (Apr. 11, 2016, 10:38 PM), http://www.bakerbotts.com/ideas/publications/2015/07/ip-report.

[146] Ryan Davis, *AIA Review Debate Now Goes to Congress*, LAW 360 (Apr. 11, 2016, 10:40 PM), http://www.law360.com/articles/676999.

[147] David O. Taylor, *Legislative Responses to Patent Assertion Entities*, 23 TEX. INTELL. PROP. L.J. 313, 314 (2015).

[148] *Id.*

[149] Davis, *supra* note 146.

[150] *Bill Summary & Status 113th Congress (2013 - 2014) H.R.3309 All Congressional Actions*, http://thomas.loc.gov (describing the Innovation Act's legislative history in the 113th Congress); *Final Vote Results for Roll Call 629*, CONGRESS.GOV, http://clerk.house.gov/evs/2013/roll629.xml (showing 325 votes in favor of the bill and 91 votes not in favor of the bill). Comment of Senator Patrick Leahy (D-Vt.), Chairman, Senate Judiciary Committee, On Patent Legislation, LEAHY.SENATE.GOV (May 21, 2014), https://www.leahy.senate.gov/press/comment-of-senator-patrick-leahy-d-vt_chairman-senate-judiciary-committee-on-patent-legislation ("Because there is not sufficient support behind any comprehensive deal, I am taking the patent bill off the Senate Judiciary Committee agenda.").

PTAB to conduct its review proceedings utilizing the claim construction that a court would use by "construing each claim of the patent in accordance with the ordinary and customary meaning of such claim as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent," rather than using the broadest reasonable construction.[152] Therefore, the Innovation Act's change in PTAB standards makes it more difficult for the PTAB to conclude invalidity because a narrower construction will be less likely to cover the prior art.

The Innovation Act's provision on claim construction standards received much support from the American Intellectual Property Law Association (AIPLA), which "currently supports, in whole or in concept, several provisions of H.R. 9, including . . . Section 9(b), which aligns the claim construction standard in *inter partes* review and post-grant review with the standard used by district courts."[153] The AIPLA, like Judge Newman, notes that this adjustment ensures that the post-grant proceedings enacted by Congress in the AIA "are being used as an effective, lower-cost alternative to challenging the validity of a patent in litigation."[154]

However, the picture is not all rosy. While the AIPLA recognizes the threat of patent trolls, the AIPLA states that it does not support many provisions of H.R. 9 as currently drafted because they "unintentionally impair the ability of patent owners to enforce their rights in good faith" and "discourage innovators who rely on the patent system for protection of their efforts."[155] Therefore, while the specific provision changing the claim construction standard may be supported, it is likely that the legislation may not move forward. Additionally, there is a question of whether Congress is the right entity to correct patent litigation problems, as this has historically been the role of courts, which have more experience on such issues.[156]

---

[151] Goodlatte Introduces Patent Litigation Reform Bill, GOODLATTE.HOUSE.GOV, (Feb. 5, 2015), http://goodlatte.house.gov/news/documentsingle.aspx?DocumentID=266.

[152] Innovation Act, H.R. 3309, 113th Cong. § 9(b)(1) (2013), https://www.gpo.gov/fdsys/pkg/BILLS-113hr3309rfs/pdf/BILLS-113hr3309rfs.pdf.

[153] Sharon Israel, *AIPLA's Views on H.R. 9, the Innovation Act of 2015*, (Apr. 16, 2015), http://www.aipla.org/advocacy/congress/114C/Documents/AIPLA%20Letter%20on%20H.R.%209%20Innovation%20Act%204-16-2015.pdf.

[154] *Id.*

[155] *Id.*

[156] *See generally* Paul R. Gugliuzza, *Patent Litigation Reform: The Courts, Congress, and the Federal Rules of Civil Procedure*, 95 B.U. L. REV. 279, 282 (2015) ("[L]egislative reform [directed to patent litigation] is unnecessary because the courts and the Judicial Conference of the United States are already in the process of fixing several problematic areas of patent litigation."). *See also* J. Jonas Anderson, *Patent Dialogue*, 92 N.C. L. REV. 1049, 1063-64 (2014)

These concerns were part of the reason Congress did not pass the Innovation Act in its initial iteration in 2014. The Act received strong criticism from judges such as Federal Circuit Judge O'Malley,[157] and from the Judicial Conference of the United States.[158] Much of the other legislation faces the same problems that plague the Innovation Act. For these reasons, Congress is not a likely vehicle for changing claim construction standards in the PTAB.

C. The USPTO Response to Lack of Amendment Opportunity

The USPTO itself has not turned a deaf ear to the criticism of the broadest reasonable interpretation. After the Federal Circuit decided *In re Cuozzo Speeding Technologies*[159] in February 2015, the USPTO organized roundtables across the country, so the public could discuss and give feedback about *inter partes* review procedures.[160] In a March 27, 2015 blog post, USPTO Director Michelle Lee announced that "as a direct result of [the] feedback" the USPTO would address the underlying criticism that it is too hard for patentees to amend claims during *inter partes* review. Lee stated that the USPTO was contemplating proposed changes to emphasize that a motion for a substitutionary amendment will always be allowed to come before the Board for consideration.[161] Under the proposed change, the issuance of amended claims would require a patent owner only to show

---

("[T]he most important changes to the patent system in recent years have been the result of the Federal Circuit *reacting* to policy signals from the Supreme Court and Congress.").

[157] Ryan Davis, *Troll Bills Would Usurp Courts' Power, Fed. Circ. Judge Says*, LAW 360 (Sept. 27, 2013), http://www.law360.com/articles/476345/troll-bills-would-usurp-courts-power-fed-circ-judge-says (indicating that Judge O'Malley said she was "stunned" that reform legislation focused on litigation case management proposals, that many of the proposals "go way beyond where anyone should want Congress to tread," and that "once you intrude on the inherent authority of courts to actually manage each case before them, you're breaking down the division between the branches of government, and there is grave danger in doing that.").

[158] H.R. REP. NO. 113-279, at 116 (2013), https://www.congress.gov/113/crpt/hrpt279/CRPT-113hrpt279.pdf (quoting a letter from the Chair of the Committee on Rules of Practice and Procedure of the Judicial Conference of the United States stating "legislation that mandates the contents of federal rules contravenes the longstanding Judicial Conference policy opposing direct amendments of the federal rules by legislation instead of through the deliberative process in the Rules Enabling Act . . . .").

[159] *In re* Cuozzo Speed Techs., LLC, 778 F.3d 1271 (Fed. Cir. 2015).

[160] Soofian & Reines, *supra* note 144.

[161] Michelle K. Lee, *PTAB'S Quick-Fixes for AIA Rules are to be Implemented Immediately, Director's Forum: A Blog from USPTO's Leadership*, USPTO (Mar. 27, 2015), http://www.uspto.gov/blog/director/entry/ptab_s_quick_fixes_for.

patentability of the narrowed amended claims beyond the prior art of record before the Office.[162]

Liberal amendments are a hallmark of the examination process, where the broadest reasonable interpretation enjoys great precedent. Therefore, if such a rule were adopted, the USPTO would have a greater justification for its use of the broadest reasonable interpretation standard in the PTAB. The Office asked for comments on the following topic: "What modifications, if any, should be made to the Board's practice regarding motions to amend?"[163] The Office received a spectrum of comments that ranged from seeking no change, to proposing liberal grant of amendments in AIA proceedings, but the USPTO maintained the status quo.[164] Instead, the USPTO believed it would be helpful to clarify the procedure for filing a motion to amend and provided an IPR, *MasterImage 3D, Inc. v. RealD, Inc.*,[165] as an example.[166] However, all that the *MasterImage* decision clarifies is that the patent owner has the burden of arguing that proposed substitute claims are patentable and narrower than the replaced claims.[167] The USPTO then, unhelpfully, stated that the burden shifts to the petitioner once the patent owner has made its *prima facie* case for patentability of the amendment.[168]

The USPTO is adamant that while there is no right to amend, there is a right to file a motion to amend. Embracing the rationale of the *In re Cuozzo Speeding Technologies* majority, the Board believes a right to file a motion is sufficient because it does not conduct a prior art search to evaluate the patentability of the proposed substitute claims.[169] The Board also feels that any such requirement would be impractical given the statutory structure of AIA proceedings.[170] Therefore, it is overly optimistic to assume that the USPTO will enact any meaningful change regarding the ability to amend.

D. Supreme Court Response

The best recourse to replace the PTAB claim constructions standard is through the Supreme Court. A couple of months after the Federal Circuit issued the

---

[162] *Id.*

[163] Amendments to the Rules of Practice for Trials Before the Patent Trial and Appeal Board, A Proposed Rule, 80 Fed. Reg. 50720 (Aug. 20, 2015) (to be codified at 37 C.F.R. pt. 42).

[164] *Id.*

[165] No. IPR2015-00040 (P.T.A.B. Jul. 15, 2015).

[166] *See* Amendments to the Rules of Practice, *supra* note 163.

[167] *MasterImage*, No. IPR2015-00040, at 1–3.

[168] *Id.*

[169] *See* Amendments to the Rules of Practice, *supra* note 163.

[170] *Id.*

In re Cuozzo en banc decision, Cuozzo asked the Supreme Court to change the claim construction standard.[171] Cuozzo appealed on the grounds that the PTAB's use of the broadest reasonable interpretation has made the *inter partes* review "surprisingly lethal" and "introduces tremendous uncertainty into claim construction [that] allows for conflicting invalidity decisions and undercuts Congress's central reform in the AIA."[172] Though noting that Congressional bills may rectify this, the possibility is low and the Supreme Court's "guidance is sorely needed on this issue critical to our patent system."[173]

Cuozzo's wish was granted. The Supreme Court approved Cuozzo's petition for writ of certiorari.[174] In the first half of 2014 alone, the Supreme Court unanimously overruled all five of the patent cases on appeal from the Federal Circuit.[175] Given the Supreme Court's track record, there is a strong chance that it could reverse the Federal Circuit's holding. A major reason the reversal rate of the Federal Circuit is so high is because the Supreme Court seems to favor fact-specific balancing-type tests over the Federal Circuit's bright-line rules.[176] Additionally, the Supreme Court's holdings and dissents on patent cases tend not to fall along any strong political lines. For example, the three recent but influential patent eligibility cases "that have thrown the industry into something of a tailspin" were all unanimous decisions.[177] Even with the passing of Justice Scalia, the

---

[171] Ryan Davis, *Cuozzo Asks Supreme Court to Overhaul AIA Review Rules*, LAW 360 (Oct. 7, 2015), http://www.law360.com/articles/711959/cuozzo-asks-supreme-court-to-overhaul-aia-review-rules.

[172] *Id.*

[173] *Id.*

[174] Cuozzo Speed Techs., LLC v. Lee, 136 S. Ct. 890 (2016).

[175] Vera Ranieri, *Supreme Court Overrules Federal Circuit Again. And Again.*, ELECTRONIC FRONTIER FOUNDATION (Jun. 2, 2014), https://www.eff.org/deeplinks/2014/06/supreme-court-overrules-federal-circuit-again-and-again.

[176] *See* KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 407, 415 (2007) (rejecting Federal Circuit's "teaching, suggesting or motivation" test as a rigid test for determining obviousness); MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 137 (2007) (rejecting Federal Circuit's bright-line rule requiring a licensee to breach or terminate a license to establish an "actual controversy" and obtain standing to pursue a declaratory judgment action); eBay Inc. v. MercExchange, L.L.C., 547 U.S. 388, 394 (2006) (rejecting Federal Circuit's categorical grant of permanent injunctions upon finding infringement and no invalidity); Illinois Tool Works, Inc. v. Indep. Ink, Inc., 547 U.S. 28, 46 (2006) (overruling Federal Circuit's presumption of market power in a patented tying product); Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 535 U.S. 722, 741-42 (2002) (rejecting Federal Circuit's interpretation of prosecution history estoppel as a complete bar to patentability).

[177] Gene Quinn, *What the Passing of Justice Antonin Scalia Means for SCOTUS Patent Jurisprudence*, IPWATCHDOG.COM (Feb. 17, 2016), http://www.ipwatchdog.com/2016/02/17/

likelihood of a split decision among the Supreme Court's eight judges is minimal. The Supreme Court's grant of certiorari should then be welcome news for Cuozzo. Furthermore, many members of the patent law community filed amicus briefs utilizing similar rationales as Judge Newman to persuade the Supreme Court to take the case.[178]

However, the Justices' commentary during oral argument suggests that change may not be as readily forthcoming. Justice Sotomayor asserted that she might be moved to agree that application of the "broadest reasonable construction" standard is inappropriate in IPR proceedings, "if Congress had not given any right for the Board to amend."[179] Justice Ginsburg noted that IPR proceedings were "[k]ind of a hybrid," "in certain respects [resembling] administrative proceedings and other district court proceedings," thereby justifying dichotomous standards.[180] Like Justice Ginsburg, Justice Kennedy pointed to no presumption of validity in IPR proceedings.[181] Justice Breyer voiced concern that "the Patent Office has been issuing billions of patents that shouldn't have been issued" insinuating that IPRs were instituted to "get rid of those patents."[182] Justice Kagan bemoaned the fact that the statute does not say one way or the other, leaving the Court "reading the tea leaves" but notices that if she were Congress and "looking at the PTO, … it does pretty much everything by this broadest-construction standard."[183]

While the aforementioned Justices expressed skepticism about removing the broadest reasonable interpretation in the PTAB, at least Justices Alito and Roberts seemed open to changing to the district court standard. Justice Alito asked if Congress had imposed upon the PTAB the same standard of proof for invalidity that is applied in the District Court, which might to indicate that perhaps Congress

---

what-justice-antonin-scalia-means-for-scotus-patent-jurisprudence/id=66247/ (referring to Mayo Collaborative Servs. v. Prometheus Labs., Inc., 132 S. Ct. 1289 (2012), Ass'n for Molecular Pathology v. Myriad Genetics, Inc., 133 S. Ct. 2107 (2013), and Alice Corp. Pty. Ltd. v. CLS Bank Int'l, 134 S. Ct. 2347 (2014)).

[178] *See*, *e.g.*, Brief of the Intellectual Property Law Association of Chicago as Amicus Curiae on Petition for Writ of Certiorari in Support of Neither Party, p. 14, Cuozzo Speed Techs., LLC v. Lee, 2016 WL 946979 (U.S.). *See also*, *e.g.*, Brief of Amicus Curiae for New York Intellectual Property Law Association in Support of Petitioner, p. 15, Cuozzo Speed Techs., LLC v. Lee, 2015 WL 7008797 (U.S.)

[179] Oral Argument at 4:10-16, Cuozzo Speed Techs., LLC v. Lee, 136 S. Ct. 890 (2016) (No. 15-446), *available at* http://www.supremecourt.gov/oral_arguments/argument_transcripts/15-446_2dp3.pdf.

[180] *Id.* at 23:21-25.

[181] *Id.* at 21:1-5.

[182] *Id.* at 14:6-15, 37:7-23.

[183] *Id.* at 21:20-25, 22:19-25.

had intended IPRs to operate under different rules than District Court litigation.[184] Justice Roberts immediately announced his position that "it seems to me perfectly clear that Congress meant for this entity to substitute for judicial action."[185] "Why," wondered Justice Roberts, "should we be wedded to the way they do business in the PTO . . . when the point is not to replicate PTO procedures."[186] He continued, "It's a very extraordinary animal in legal culture to have two different proceedings addressing the same question that lead to different results."[187]

While a minority of justices outwardly expressed support for *In re Cuozzo*'s dissent, to allay the fears of practitioners and provide a stable environment with which the IPR claim construction standard can progress, the Supreme Court is still the best vehicle for eliminating the PTAB's broadest reasonable interpretation. The Supreme Court not only bypasses the USPTO's intransigence on easing up the ability to amend but also provides a quicker path to rectifying the disparity in standards than Congress has done thus far. However, if the Supreme Court does not reverse the Federal Circuit decision, then it will be up to Congress to swiftly provide a solution to the dilemma of the broadest reasonable interpretation in the PTAB.

CONCLUSION

Claim construction is perhaps the most important issue in a determination of invalidity for patents. Thus, there is a quagmire in the dichotomy of patent claim constructions that exists between the PTAB and district courts. It is well known that the USPTO adopted the broadest reasonable interpretation, for at least the past hundred years, to ascertain the upper bound of claim meaning during patent examination. The broadest reasonable interpretation standard was then adopted for the new post-grant proceedings created by the AIA. While cost and time efficiency are partly to explain IPRs sharp rise to popularity, the broadest reasonable interpretation also plays a large part in informing petitioners' preferences for the IPR over district court litigation. As the claim construction is broader, the chance of invalidating the patent owner's claims increases.

The disparity between the district court's "ordinary and customary meaning" claim construction standard and the USPTO's broadest reasonable interpretation standard creates substantial risk for patent owners. The disparity also promotes inefficient litigation when a district court does not grant a stay pending an IPR

[184] *Id.* at 6:23-7:1.
[185] *Id.* at 31:4-6.
[186] *Id.* at 31:7-11.
[187] *Id.* at 32:2-5.

proceeding. Because there is a disparity across district courts in granting stays, the USPTO's claim interpretation standard contributes to patent litigation forum shopping. The Federal Circuit's *In re Cuozzo Speeding Technologies* case highlights many of the pitfalls that surround the PTAB's use of the broadest reasonable interpretation for post-grant proceedings. Judge Newman's dissent noted how the majority's holding undermined the very purpose of the *inter partes* review: "providing quick and cost effective alternatives to litigation"[188] in order to provide "improved service to technology-based innovation [and] . . . creative advance and industrial growth."[189]

The 6-5 Federal Circuit split in *In re Cuozzo Speeding Technologies* was a clear indication that action needs to be taken to address the problem that afflicts the PTAB claim construction standard. There are several avenues for recourse but only the Supreme Court has any promise of achieving meaningful change. In the pending appeal, the Supreme Court should overrule the majority decision in *In re Cuozzo Speeding Technologies* and adopt Judge Newman's view that the broadest reasonable interpretation was never intended to be utilized in the PTAB, especially with such minimal ability to amend claims.

By using the district court standard in the PTAB, as this note advocates, IPRs will still maintain their time and cost efficiency and remain the popular alternative to district court litigation, as Congress intended. However, aside from the maintenance of popularity, there will also be a greater benefit served by removing the broadest reasonable interpretation in the PTAB. The comment process conducted by the USPTO, as discussed in Part III, showed that the USPTO would not budge on changing their amendment practice.[190] Switching to the district court claim construction standard would allow the PTAB to maintain its stringent IPR amendment practice because the broadest reasonable interpretation is only justified where amendment may be made liberally. When the broadest reasonable interpretation does not exist, a right to amend may be limited. Moreover, district courts would no longer have a reason to deny staying litigation in the midst of a pending IPR because the definitions of a patent claim would not be disparate. Litigating in two different venues, the PTAB and district courts, will no longer yield the possibility of opposing validity holdings. With a unification of claim construction standards, if a patent is valid in a PTAB proceeding, it must be valid in a district court.

---

[188] H.R. REP. NO. 112-98, pt. 1, at 48 (2011).
[189] *In re Cuozzo*, 793 F.3d at 1287.
[190] *See supra* Part III.C.

The problems that accompany a disparity of claim construction standards in the PTAB and district courts make patent owners skeptical of their future patents. They dread the possibility of litigating on two fronts under two separate claim construction standards. For all the aforementioned benefits that would proceed from a unification of standards in the PTAB and district courts, the Supreme Court should alleviate this disparity by eliminating the broadest reasonable interpretation from the PTAB.

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 5 SPRING 2016 NUMBER 2

ANTITRUST TREATMENT OF THE NO CHALLENGE CLAUSE

THOMAS K. CHENG*

This Article examines a patent licensing practice that has hitherto escaped the attention of U.S. antitrust law: the no challenge clauses. Under these clauses, a patent licensee is prohibited from challenging the validity of the licensed patent. These clauses have so far only been examined under patent law in terms of enforceability. This oversight by antitrust law is unfortunate, as no challenge clauses can create consumer harm by protecting an otherwise invalid patent from challenges and artificially extending the exclusive period granted by the patent law. This means that consumers have to bear supra-competitive prices for longer than necessary. A number of factors are relevant to the analysis of the legality of no challenge clauses, such as market power, patent validity, and market structure at the licensee level. This Article proposes a framework based on the Rule of Reason that incorporates all of these relevant factors and structures them in a way that renders the framework easy to apply. Lastly, the Article rejects a number of justifications that have been offered to argue for the legality of these clauses across the board.

* Associate Professor, Faculty of Law, University of Hong Kong; thomas.cheng@hku.hk. The Author would like to thank Joshua Kanjanapas, Allison Wong, and Joel Lee for their able research assistance.

INTRODUCTION

One of the anomalies in U.S. antitrust law, despite its comparatively vast jurisprudence, is that courts seem to have never ruled on the legality of no challenge clauses. These clauses generally prohibit a licensee from challenging the validity of the patent being licensed. Thus far, no challenge clauses have only attracted the attention of patent law and have been dealt with largely as a question of enforceability.[1] Different appellate courts have expressed diverse views on the enforceability of no challenge clauses, depending on the nature and timing of the agreement in which they are incorporated.[2] To the extent that these clauses are enforceable, it leads to the question of how they should be treated under antitrust law.

In the circuits in which they are unenforceable, no challenge clauses may be viewed as a non-binding agreement by licensees that they will not challenge the validity of the patent. In such cases, there is a legitimate question as to why licensees would make such a commitment, what enticements have been offered by the patentee to secure such a commitment, and what this tells us about the patentee's own belief in the likelihood of patent validity.

In circuits in which no challenge clauses are enforceable, these clauses can exert anticompetitive effects by preventing challenges to invalid patents. No challenge clauses do not seem so competitively benign that one can conclusively assert that they do not inflict harm on consumers.

A focus on no challenge clauses is further justified by the attention that other jurisdictions have paid to them in recent years. In 2015, a Chinese enforcement authority fined Qualcomm close to $1 billion over the imposition of no challenge clauses, among other offenses.[3] Moreover, no challenge clauses are one of the areas in which the U.S. and the European Union ("EU"), the two leading antitrust jurisdictions in the world, have diverged. While U.S. antitrust law has largely left these clauses untouched, the EU, to the extent its view is embodied by the

[1] *See* discussion *infra* Section II.A.
[2] *See* discussion *infra* Section II.A.
[3] *See* discussion *infra* Section II.C.2.

European Commission, has taken a fairly hostile attitude toward them.[4] In fact, largely due to the Commission's view, these clauses have mostly been expunged from European licensing agreements.[5] In light of the international divergences, it is worth exploring how these clauses should be approached.

This Article fills an important gap in the U.S. antitrust academic literature by exploring antitrust treatment of no challenge clauses. As far as this author is aware, no academic article thus far has examined this issue. The only notable exception is an article by Miller and Gal, which focused on the enforceability of no challenge clauses from a patent law and total welfare perspective.[6]

This Article is divided into seven sections. Following this introductory section, Section I provides an overview of no challenge clauses and sets forth a taxonomy for these clauses. Section II surveys the jurisprudence on no challenge clauses in the three main antitrust jurisdictions in the world; the U.S., the EU, and China, highlighting the differences among them. Section III explains the circumstances under which no challenge clauses can create consumer harm and identifies the relevant factors for analyzing and predicting such harm. Section IV enumerates the various justifications for no challenge clauses and rebuts them. Section V summarizes the main ideas in the preceding Sections and outlines an analytical framework for analyzing no challenge clauses under antitrust law. The conclusion is the final section.

I
OVERVIEW OF NO CHALLENGE CLAUSES

A. Definition of No Challenge Clauses

No challenge clauses are inserted in patent licensing agreements to prohibit the licensee from challenging the validity of the patent for a period of time, usually the duration of the contract.[7] Patentees incorporate such clauses into their licensing agreements to forestall potential validity challenges by the licensees. According to Orstavik, "[t]he object of a no-challenge clause is to fortify a position granted by

[4] *See* discussion *infra* Section II.A.2.

[5] *See* Sophie Lawrence, *The Competition Law Treatment of No-Challenge Clauses in License Agreements: An Unfortunate Revolution?*, 9(10) J. INTELL. PROP. L & PRAC. 802, 810 (2014).

[6] Alan D. Miller & Michal S. Gal, *Licensee Patent Challenges*, 32 YALE J. ON REG. 122, 127 (2015).

[7] Inger B. Orstavik, *Technology Transfer Agreements: Grantbacks and No Challenge Clauses in the New EC Technology Transfer Regulation*, 36(1) INT'L REV. INTELL. PROP. & COMPETITION L. 83, 87 (2005).

law."[8] These clauses, however, do not provide patentees with fool proof defenses against validity challenges; because they only govern licensees, they have no effect on the conduct of unrelated third parties.[9] Therefore, the patent could still be subject to challenges by third party actors. The degree of protection offered by these clauses therefore depends on the likelihood and willingness of unrelated third parties to challenge the patent. If there is a third party that is likely and willing to challenge the patent despite the clause, the degree of protection afforded to the patentee will be limited. However, if most of the possible challengers have already been recruited as licensees and are subject to the no challenge clause, the patentee can be assured of the continual validity of its patent.

In spite of the no challenge clause, the likelihood of third party challenges to the patent bears upon the continual validity of the patent and its competitive effects. Ultimately, this likelihood is circumstance-specific and requires detailed examination. The courts and commentators, however, have opined that licensees are the parties with the greatest economic incentives to challenge the validity of patents. In *Lear, Inc. v. Adkins*, the U.S. Supreme Court declared that "[l]icensees may often be the only individuals with enough economic incentive to challenge the patentability of an inventor's discovery."[10] To the extent that this is true, no challenge clauses will effectively forestall validity challenges, which may allow an invalid patent to persist.

B. Different Types of No Challenge Clauses

There are different types of no challenge clauses, which can be classified into two main categories. The first category consists of outright prohibitions of validity challenges in the licensing agreement, which Miller and Gal have called no contest clauses.[11] Whether the clause in fact achieves outright prohibition depends on the willingness of the courts to grant injunctions or to bar validity challenges to enforce these clauses. To the extent that courts eschew injunctions for enforcing no challenge clauses, the patentee will only obtain damages. In that case, outright validity challenge prohibitions will only impose a financial penalty, which places them in the second category. Miller and Gal call these challenge penalty clauses.[12]

There is a wide variety of challenge penalty clauses. The penalty may be in the form of a financial penalty or a loss of contractual privileges, which ultimately

---

[8] *Id.*
[9] *Id.*
[10] Lear, Inc. v. Adkins, 395 U.S. 653, 670 (1969).
[11] Miller & Gal, *supra* note 6, at 127.
[12] *Id.* at 131.

will result in financial losses for the licensees. The financial penalty can be in the form of liquidated damages or higher royalties. For instance, in *Rates Technology Inc. v. Speakeasy, Inc., LLC*, the no challenge clause stipulated liquidated damages of a value of over twenty-four times the license fee.[13] One might argue that if the liquidated damages are so substantial that it would have a serious financial impact on the licensee, or perhaps even bankrupt the licensee, the challenge penalty clause effectively functions as an outright prohibition. The financial penalty may also exist in the form of elevated royalty.[14] In such case, the challenge penalty clause would stipulate that the royalty rate would increase in response to a validity challenge launched by the licensee.[15] A slight variation of an elevated royalty clause is a clause that provides for a higher royalty rate only when the validity challenge turns out to be unsuccessful.[16] A further variation is a clause that establishes three tiers of royalty rates, "with the rate increasing once a challenge is mounted, and providing for an even higher royalty if the challenge is not successful."[17] These various types of clause create financial disincentives for the licensees to challenge a patent.[18]

Another kind of arrangement that similarly creates financial disincentives for licensees to challenge a patent is royalty front-loading.[19] Strictly speaking, this type

---

[13] Rates Tech., Inc. v. Speakeasy, Inc., 685 F.3d 163 (2d Cir. 2012).

[14] Lorelei Ritchie, *Reconciling Contract Doctrine with Intellectual Property Law: An Interdisciplinary Solution*, 25 SANTA CLARA COMPUTER & HIGH TECH. L.J. 105, 146 (2008).

[15] *Id*.

[16] Rochelle Cooper Dreyfuss & Lawrence S. Pope, *Dethroning Lear? Incentives to Innovate After MedImmune*, 24 BERKELEY TECH L.J. 971, 1001 (2009).

[17] *Id*.

[18] To the extent that the royalty increase kicks in only after the patent has been validated, Rochelle Dreyfuss and Lawrence Pope argue that the clause does not impose a penalty at all; the royalty increase merely reflects the increased value of a patent that has survived a challenge. *Id*. at 1002. There is a general perception that a patent that has been validated by the courts is more valuable than an untested patent. While a validated patent is no doubt more valuable to the patentee, it is not entirely clear why the patent would become more valuable to the licensee. To the licensee, a license is valuable because it allows the licensee to use the patented technology. This right to use the patented technology should not change in value after validation. Validated or not, what is valuable to the licensee is not the right to exclude granted by the patent, but the underlying technology, which does not change after the patent has been validated. A license to an invalidated patent will be worth less (or perhaps nothing) because everyone is now free to use the technology. But a license to a validated and an invalidated (a patent that has not been subject to a validity challenge) patent should be worth the same to the licensee. The only way in which a license to a validated patent may be worth more to a licensee is if the current market is not entirely competitive and the licensee is able to charge a somewhat supra-competitor price, and a validated patent will be able to exclude third parties without a license with certainty.

[19] Lawrence, *supra* note 5, at 810.

of arrangement need not entail a financial penalty. If the royalty that is required of the licensee over the duration of the licensing agreement is the same as the amount that the licensee is liable to pay without front-loading, there is no financial penalty on the licensee. However, the licensee would be similarly deterred from challenging the patent as compared to a royalty increase upon challenge. This is because under current patent law, a licensee cannot recover the royalty that has been paid to the patentee prior to invalidation if the patent turns out to be invalidated, even though the licensee arguably should have never had to pay royalty to the patentee if the patent had always been invalid.[20] Therefore, if a patentee front-loads the royalty, the licensee will lose the incentive to challenge the patent because she will achieve little savings in terms of aggregate royalty payment. Even though the arrangement does not entail a financial penalty, it would achieve a similar result as a challenge penalty clause.[21]

The remaining type of challenge penalty clause is the termination-upon-challenge clause, which stipulates a termination of the licensing agreement upon the launch of a validity challenge by a licensee. This type of clause functions mainly by threatening the licensee with damages claims from the patentee if the licensee continues to deploy the licensed technology to produce the product. Upon the launch of a validity challenge, the licensing agreement either automatically terminates or gives the patentee an option to terminate the agreement.[22] Once the agreement is terminated, the licensee would be infringing the patent if she chose to continue to use the technology. If the patent is eventually upheld, the patentee can sue the licensee for patent infringement. The licensee may even be liable for trebled damages if the patentee can prove that the infringement is willful.[23] This gives the licensee a significant disincentive to bring validity challenges, at least unless she is quite confident of her chance of success. This may serve the laudable purpose of deterring frivolous validity challenges,[24] but the deterrent effect may be so great that it discourages meritorious challenges that are short of a slam-dunk.

---

[20] See Nellie A. Fisher, *The Licensee's Choice: Mechanics of Successfully Challenging a Patent under License*, 6 TEX. INTELL. PROP. L.J. 1, 31-43 (1997).

[21] Dreyfuss & Pope, *supra* note 16, at 994. Commentators, however, have noted the limitations of royalty front-loading. In particular, it has been argued that front-loading may not be feasible if the licensee is cash strapped or if the commercialization of the technology requires substantial upfront investment. *See id.* at 983, 992-996; Miller & Gal, *supra* note 6, at 150. In that case, the licensee may be unable or unwilling to pay a substantial part of the royalty upfront.

[22] Christian Chadd Taylor, *No-Challenge Termination Clauses: Incorporating Innovation Policy and Risk Allocation into Patent Licensing Law*, 69 IND. L.J. 215, 230 (1993).

[23] Patent Act § 284, 35 U.S.C. § 284 (2011).

[24] Taylor, *supra* note 22, at 234 (1993).

Because of the mixed effects of termination-upon-challenge clauses, there are differing views as to whether they actually impose a penalty. Some commentators have argued that termination-upon-challenge clauses merely level the playing field between the patentee and the licensee in the course of litigation and will help promote innovation by protecting the patentee's investment.[25] According to Taylor, "[d]uring litigation, the licensee profits from the product without paying royalties or incurring competition from other licensees. The licensor, on the other hand, must incur litigation without collecting royalties and, if the licensee holds an exclusive license, without the right to license the patent to another."[26] Short of repudiating the licensing agreement, the licensee could continue to produce the product using the patentee's technology while challenging the patent and holding the patentee bound by the agreement. Some critics argue that this is unfair to the patentee.[27] Meanwhile, other commentators have contended that termination-upon-challenges may have a deleterious effect on welfare and should be subject to scrutiny by the courts.[28] Regardless of whether the patentee is in a disadvantageous bargaining position vis-à-vis the licensee in the course of a validity challenge, it is clear that termination-upon-challenge clauses produce significant deterrent effect on licensees. Especially if willful infringement can be proved, the effect of these clauses could be similar to that of no challenge clauses with hefty liquidated damages, as in *Rates Technology*.[29]

The treatment of these various types of outright no challenge and challenge penalty clauses under U.S. patent law is still subject to debate. While most believe that an outright no challenge clause in a licensing agreement would be unenforceable, there is case law that suggests otherwise.[30] The situation is likewise unclear for termination-upon-challenge clauses.[31] Furthermore, while some commentators believe that the Supreme Court would invalidate no challenge clauses after *MedImmune, Inc. v. Genentech, Inc*,[32] others believe that the issue is still wide open.[33] With respect to the various royalty adjustment mechanisms, given

---

[25] *Id*. at 232.
[26] *Id*. at 243.
[27] *Id.*
[28] Miller & Gal, *supra* note 6, at 154.
[29] *See generally* Rates Tech., Inc. v. Speakeasy, Inc., 685 F.3d 163 (2d Cir. 2012).
[30] *See* discussion *infra* Section II.A.2.
[31] *See* discussion *infra* Section II.A.5.
[32] See M. Natalie Alfaro, *Barring Validity Challenges Through No-Challenge Clauses and Consent Judgments: MedImmune's Revival of the Lear Progeny*, 45 HOUS. L. REV. 1277, 1309 (2008).
[33] *See* Dreyfuss & Pope, *supra* note 16, at 1004-05. ("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

the latitude that is usually given to the patentees to structure their royalty, it is unlikely that they will be deemed unenforceable.

Nonetheless, enforceability under patent law and legality under the antitrust law are two distinct issues. When determining whether a certain licensing practice should be enforceable under patent law, courts usually look to patent policy, which aims to encourage innovation by generating sufficient incentives.[34] When assessing the legality of the same practice under antitrust law, courts pay heed to antitrust policy and the overriding objective of the protection of consumer welfare.[35] Antitrust law emphasizes substance over form. If an outright no challenge clause and the various challenge penalty clauses exert the same effect on the licensee's incentive to launch a validity challenge, they should result in similar impact on consumer welfare and therefore should be accorded the same treatment. The form in which the deterrent effect on licensees is achieved should not be dispositive from an antitrust perspective.

C. Timing of the Agreement

Apart from the form in which they take, no challenge clauses also may differ in the timing between when the agreement is entered and when the validity challenge commences and concludes. There are generally three time settings for the entry of the agreement. The first is a pure licensing agreement that is entered in the absence of any threat of litigation. The second is a settlement agreement that is entered when litigation is imminent or has commenced and has proceeded to various stages prior to conclusion. The third is consent decree, which concludes litigation by the agreement of both parties with the court's approval. The question, therefore, is whether the timing of the agreement affects the enforceability of the no challenge clause under patent law and should affect the legality of the clause under antitrust law.

Overall, the timing of the agreement has had a bearing on judicial attitude toward no challenge clauses, although there is no clear consensus among the appellate courts.[36] Courts seem to have treated no challenge clauses in licensing agreement with the greatest hostility.[37] Most seem to agree that no challenge

---

licensee choose to challenge the validity of the patent. With respect to litigation risks, this would fully restore the parties to the pre-Medlmmune situation.").

[34] HERBERT HOVENKAMP ET AL., IP AND ANTITRUST: AN ANALYSIS OF ANTITRUST PRINCIPLES APPLIED TO INTELLECTUAL PROPERTY LAW § 1.3.

[35] *Id.*

[36] *See* discussion *infra* Section II.A.

[37] *See* discussion *infra* Section II.A.2.

clauses incorporated in licensing agreements are unenforceable as a matter of patent law.[38] However, the Supreme Court has never expressly decided the issue. The Federal Circuit, which is the most important appellate court for patent issues, has indicated that a clear and unambiguous no challenge clause should be enforceable even in the absence of a threat of litigation.[39]

No challenge clauses in settlement agreement seem to have received more lenient treatment from the courts.[40] Again, consensus eludes the various appellate courts. The Second and the Ninth Circuits have struck down no challenge clauses in settlement agreements,[41] while the Sixth and the Federal Circuits have upheld them under specific circumstances.[42] More recently, the Second Circuit has held that a no challenge clause contained in a settlement agreement entered into after discovery would be upheld.[43] A similar circuit split is also observed with respect to the enforceability of no challenge clauses incorporated in consent decrees. The Second and the Seventh Circuits have refused to enforce no challenge clauses contained in consent decrees[44] whereas the Federal Circuit has largely enforced them.[45]

Courts have offered a range of reasons for offering disparate treatment to no challenge clauses contained in different types of agreements. For consent decrees, the Federal Circuit has argued that the doctrine of res judicata favors the definitive

---

[38] *See* discussion *infra* Section II.A.2.

[39] Baseload Energy, Inc. v. Roberts, 619 F.3d 1357, 1363 (Fed. Cir. 2010) ("In the context of settlement agreements, as with consent decrees, clear and unambiguous language barring the right to challenge patent validity in future infringement actions is sufficient, even if invalidity claims had not been previously at issue and had not been actually litigated.").

[40] *See* discussion *infra* Section II.A.3.

[41] Massillon-Cleveland-Akron Sign Co. v. Golden State Advert. Co., 444 F.2d 425 (9th Cir. 1971) (striking down no challenge clauses in the 9th Circuit); Warner-Jenkinson Co. v. Allied Chemical Corp., 567 F.2d 184 (2d Cir. 1977) (striking down no challenge clauses in the 2nd Circuit).

[42] Aro Corp. v. Allied Witan Co., 531 F.2d 1368 (6th Cir. 1976) (enforcing no challenge clause in settlement agreement entered into after discovery); Flex-Foot, Inc. v. CRP, Inc., 238 F.3d 1362 (Fed. Cir. 2001) (enforcing no challenge clause where alleged infringer had challenged patent validity, had had opportunity to conduct discovery regarding validity, and had agreed voluntarily to dismiss suit with prejudice).

[43] Rates Tech., Inc. v. Speakeasy, Inc., 685 F.3d 163 (2d Cir. 2012) (enforcing no challenge clause in settlement agreement entered into after discovery).

[44] Addressograph-Multigraph Corp. v. Cooper, 156 F.2d 483 (2d Cir. 1946) (refusing to enforce no challenge clause in consent decrees); Bus. Forms Finishing Serv. v. Carson, 452 F.2d 70 (7th Cir. 1971).

[45] Foster v. Hallco Mfg., 947 F.2d 469 (Fed. Cir. 1991); Diversey Lever, Inc. v. Ecolab, Inc., 191 F.3d 1350 (Fed. Cir. 1999).

disposition of legal disputes, and parties should not be allowed to reopen the validity issue later.[46] For settlement agreements, there seems to be a predominant view that discovery affords parties the opportunity to fully inform themselves of the issues. Thus a no challenge clause incorporated in a settlement agreement entered into after discovery should represent an informed, binding decision by the parties.[47]

Whether no challenge clauses contained in licensing and settlement agreements should be treated differently from an antitrust perspective will be discussed subsequently. For now, suffice it note that consent decrees present slightly different issues from the other two types of agreements. While some courts have admittedly refused to enforce no challenge clauses contained in consent decrees, attaching antitrust liability to the clause is a different matter.[48] Even though the basis of a consent decree is an agreement between the two litigating parties, judicial supervision would suggest that the court approves of the provisions in the agreement.[49] It would be quite remarkable to assert that a clause that has been approved by the court should turn out to be illegal, giving rise to trebled damages and other liability. Therefore, no challenge clauses contained in consent decrees should be beyond the purview of antitrust law. Instead, the focus should be on licensing and settlement agreements.

II
TREATMENT OF NO CHALLENGE CLAUSES IN DIFFERENT JURISDICTIONS

The treatment of no challenge clauses varies widely across the major jurisdictions. In the U.S., no court seems to have ruled on the legality of no challenge clauses under antitrust law or held them to constitute patent misuse.[50] As

---

[46] *Foster*, 947 F.2d at 474-75.

[47] Dylan Pittman, *Allowing Patent Validity Challenges Despite No-Challenge Clauses: Fulfilling the Will of King Lear*, 48 IND. L. REV. 339, 356 (2014); Melissa Brenner, Comment, *Slowing the Rates of Innovation: How the Second Circuit's Ban on No-Challenge Clauses in Pre-Litigation Settlement Agreements Hinders Business Growth*, 54 B.C. L. REV. SUPP. 57 (2013).

[48] *See, e.g.*, *Addressograph-Multigraph*, 156 F.2d 483; *Bus. Forms Finishing Serv.*, 452 F.2d 70.

[49] Judith Resnik, *Judging Consent*, 1 U. CHI. LEGAL F. 43, 45 (1987).

[50] In *Bendix*, the Seventh Circuit did rule on the issue of whether a no challenge clause can constitute patent misuse during and after the term of the license. Bendix Corp. v. Balax, Inc., 421 F.2d 809 (7th Cir. 1970). It held that such a clause during the term of the license did not constitute patent misuse. The Court decided the case largely on the policy articulated in *Lear, Inc. v. Adkins*, and did not consider antitrust policy. The Court did hold that a post-expiration no challenge clause may constitute a patent misuse under *Brulotte v. Thys Co.*, which had held that

mentioned earlier, the various appellate courts have expressed different views on their enforceability under patent law. In the EU, the position on no challenge clauses under the Technology Transfer Block Exemption Regulations ("TTBER") has evolved over time. The current position under the 2014 TTBER is that all no challenge clauses, including termination-upon-challenge clauses, fall within what are known as excluded restrictions.[51] Apart from a limited number of exceptions, these clauses will not benefit from the block exemption and will need to be justified under Article 101(3) of the Treaty on the Functioning of the European Union ("TFEU") in order to be lawful under EU competition law. For all intents and purposes, most parties avoid clauses that are excluded restrictions in their licensing agreements, partly because justification under Article 101(3) is generally perceived to be difficult.[52] In other words, commercial parties practically treat no challenge clauses as illegal.

In China, one of the relatively recent but nonetheless important jurisdictions, no challenge clauses also seem to be practically illegal per se, as indicated by the one case in which they were examined. In the February 2015 decision on Qualcomm's licensing practices, the National Development and Reform Commission ("NDRC), one of the Chinese enforcement authorities, effectively deemed no challenge clauses as illegal per se. In the IP-Competition Regulations issued in April 2015 by the State Administration of Industry and Commerce ("SAIC") (another Chinese enforcement authority), Article 10 proscribes the use of no challenge clauses absent legitimate justifications.[53] The Regulations are silent on what constitutes a legitimate justification. The following sections provide a detailed overview of the status of no challenge clauses under U.S., EU, and Chinese law.

---

collection of royalty post expiration is patent misuse. Again, the decision was not made on antitrust grounds. A number of courts have reached a similar conclusion regarding pre-expiration no challenge clauses. *See* Congoleum Ind., Inc. v. Armstrong Cork Co., 366 F. Supp. 220 (E.D. Penn. 1973); Wallace Clark & Co., Inc. v. Acheson Ind., Inc., 401 F. Supp. 637 (S.D.N.Y. 1975); Panther Pumps & Equipment. Co. v. Hydrocraft, Inc., 468 F.2d 225 (7th Cir. 1972).

[51] Commission Regulation (EU) 316/2014 of Mar. 21, 2014, The Application of Article 101(3) of the Treaty on the Functioning of the European Union to Categories of Technology Transfer Agreements, OJ L93/17 [hereinafter "2014 TTBER"], art. 5(1)(b).

[52] Lawrence, *supra* note 5, at 2.

[53] Guanyu Jinzhi Lanyong Zhishi Chanquan Paichu, Xianzhi Jingzheng Xingwei de Guiding (关于禁止滥用知识产权排除、限制竞争行为的规定) [Provisions on Prohibiting the Abuse of Intellectual Property Rights to Exclude and Restrain Competition] (promulgated by State Administration of Trade and Commerce, Apr. 7, 2015, effective Aug. 1, 2015), http://www.saic.gov.cn/zcfg/xzgzjgfxwj/xxb/201504/t20150413_155104.html.

A. The United States

1. Lear, Inc. v. Adkins

U.S. courts have suggested that no challenge clauses would be illegal if they were incorporated into a market allocation agreement.[54] It has also been held that the use of reciprocal dealing to force a counterparty not to challenge a patent is an antitrust violation.[55] However, it seems that no courts have ruled on the legality of no challenge clauses on their own; instead, much of the action regarding no challenge clauses has been under patent law. Any exposition of the law on no challenge clauses must start with the 1969 Supreme Court case of *Lear, Inc. v. Adkins*. Prior to this case, the doctrine of licensee estoppel, which was first applied by the Supreme Court in 1856 in *Kinsman v. Parkhurst*,[56] had prevailed in the U.S.[57] The doctrine essentially states that once a licensee accepts a licensing agreement from a patentee, the licensee is deemed to have acquiesced to the

---

[54] In *Jack Winter, Inc.*, the Court held that a mere agreement not to challenge the validity of a patent without an accompanying market division agreement does not constitute an illegal per se market allocation agreement under the Sherman Act. Jack Winter, Inc. v. Koratron Co., Inc., 375 F. Supp. 1 (N.D. Cal. 1974). The Court also held that the agreement was not an unreasonable restraint of trade based on a variety of reasons that did not focus on the competitive harm of the agreement.

In *Nachman Spring-Filled Corp*., the Court held that a clause in an agreement whereby a party acknowledged the validity of a patent is illegal under the Sherman Act. Nachman Spring-Filled Corp. v. Kay Mfg. Co., 139 F.2d 781 (2d Cir. 1943). However, the agreement at issue also contained a market allocation agreement whereby one party agreed to cease production. And the Court's holding that the validity acknowledgement clause is illegal is closely tied to the legality of the market allocation agreement. "Accordingly defendant's covenant acknowledging the patent's validity constitutes, in effect, an undertaking that, if sued by plaintiff for enforcement of that agreement, defendant will not assert the defense that the agreement is illegal. Such a raising-by-one's-boot's-straps undertaking, of course, cannot be enforced." *Id*. at 784.

[55] In *W.L. Gore & Assoc.*, the Court held that the use of reciprocal dealing to coerce an alleged infringer to not challenge the validity of a patent to be patent misuse and an antitrust violation. W.L. Gore & Assoc. v. Carlisle Corp., 381 F. Supp. 220 (E.D. Penn. 1973). However, the Court's emphasis was clearly on the infringement plaintiff's use of reciprocal dealing, and not the no challenge clause per se. Moreover, there was no agreement of any kind between the two parties. The case merely concerned a threat against the other party not to challenge the patent. *Id*.

[56] Kinsman v. Parkhurst, 59 U.S. 289 (1855).

[57] However, as the Court itself noted in *Lear*, the Court had never consistently applied the doctrine since *Parkhurst*. In a few subsequent decisions, the Court refused to apply the doctrine to estop licensee validity challenge without much effort to distinguish the instant case from *Parkhurst*. In the first half of the 20th century, the Court had created so many exceptions to the doctrine that "the estoppel doctrine had been so eroded that it could no longer be considered the 'general rule'". Lear Inc. v. Adkins, 395 U.S. 653, 664 (1969).

validity of the patent underlying the agreement and is estopped from launching validity challenges later.[58] The doctrine was largely based on equitable considerations and paid little heed to the social harm of upholding an invalid patent.[59]

In *Lear, Inc. v. Adkins*, the U.S. Supreme Court ruled on whether the licensee estoppel doctrine estopped Lear, Inc. from pleading patent invalidity in the suit. In language that has been cited repeatedly by the lower courts ever since, the Supreme Court declared that the public policy of clearing invalid patents overrides the equitable considerations favoring the patentee:

> Surely the equities of the licensor do not weigh very heavily when they are balanced against the important public interest in permitting full and free competition in the use of ideas which are in reality a part of the public domain. 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 in the typical situation involving the negotiation of a license after a patent has issued.[60]

This paragraph is notable for developing the law on no challenge clauses in two respects. First, although the case did not in fact involve a no challenge clause,[61]

---

[58] ABA SECTION OF ANTITRUST LAW, INTELLECTUAL PROPERTY AND ANTITRUST HANDBOOK 233 (2007).

[59] *See Lear*, 395 U.S. at 669-70 (discussing whether the doctrine applies based on equities of licensor).

[60] *Id.* at 670-71.

[61] Miller and Gal, however, argue that one of the contractual provisions in the case effectively functioned as a no challenge clause because it "required the licensee to continue paying royalties during the pendency of the patent challenge." Miller & Gal, *supra* note 6, at 131. It is unclear whether the practical effect of this clause is such that it functions as a no challenge clause. The clause effectively reduces the payoff to the licensee for a successful challenge by the amount of royalty due during the litigation. Assuming that the patent is not nearing expiration (in which case the licensee would have few incentives to challenge the patent anyway), and the ratio between the litigation period and the remainder of the patent term (assuming that the licensee intends to renew the licensing agreement all the way up to patent expiration) is not very high, there is no reason to believe that the reduction in payoff should have a significant effect on the licensee's incentive to mount a validity challenge. Most other commentators tend to agree that *Lear* did not concern a no challenge clause. *E.g.*, Taylor, *supra*

lower courts have cited the balance of public policy in favor of the removal of invalid patents as justification for invalidating no challenge clauses of various kinds.[62] Some commentators have argued that *Lear* does not require this result at all.[63] Second, subsequent courts and commentators alike have cited with approval the court's observation that licensees are often the only parties with economic incentives to mount a validity challenge.[64] The circumstances that affect a licensee's incentive to challenge will be discussed subsequently.

Appellate courts applied Lear's holdings to no challenge clauses in the ensuing decades.[65] There is quite a divide between the courts on their treatment of no challenge clauses. By and large, the Federal Circuit, unsurprisingly, has taken a pro-patentee approach and allowed these clauses to be enforced under various

---

note 22, at 231; Brenner, *supra* note 47, at 62; Alfaro, *supra* note 32, at 1286. Miller and Gal, however, argue that one of the contractual provisions in the case effectively functioned as a no challenge clause because it "required the licensee to continue paying royalties during the pendency of the patent challenge." Miller & Gal, *supra* note 6, at 131. It is unclear whether the practical effect of this clause is such that it functions as a no challenge clause. The clause effectively reduces the payoff to the licensee for a successful challenge by the amount of royalty due during the litigation. Assuming that the patent is not nearing expiration (in which case the licensee would have few incentives to challenge the patent anyway), and the ratio between the litigation period and the remainder of the patent term (assuming that the licensee intends to renew the licensing agreement all the way up to patent expiration) is not very high, there is no reason to believe that the reduction in payoff should have a significant effect on the licensee's incentive to mount a validity challenge. Most other commentators tend to agree that *Lear* did not concern a no challenge clause. *E.g.*, Taylor, *supra* note 22, at 231; Brenner, *supra* note 47, at 62; Alfaro, *supra* note 32, at 1286.

[62] *See, e.g.*, Massillon-Cleveland-Akron Sign, 444 F.2d at 428 ("If a patent holder can exact from another a promise not to infringe, and thereby recover from one inducing the breach of that promise, in the absence of a valid patent, the patent holder is afforded more protection than the patent laws allow. The patent holder acquires this additional protection 'merely because he (MCA here) chose one remedy (inducement to breach a contract not to infringe) rather than another (inducement to infringe) on the same substantive issue.' Federal policy favoring free competition in ideas not meriting patent protection cannot be so easily subverted.") (internal citations omitted); Bendix Corp. v. Balax, Inc., 421 F.2d 809, 821(7th Cir. 1970) ("From all this we can only conclude that the right to estop licensees from challenging a patent is not part of the 'limited protection' afforded by the patent monopoly.").

[63] Taylor, *supra* note 22, at 231; Brenner, *supra* note 47, at 62; Alfaro, *supra* note 32, at 1286.

[64] *Bendix*, 421 F.2d at 809; *Rates Tech. v. Speakeasy, Inc.*, 685 F.3d 163 (2d Cir. 2012); Miller & Gal, supra note 6, at 137 ("Patent licensees are in a special position to perform this role. Their practical experience with the subject matter of the patent often places them in a good position to evaluate the novelty of the invention. They might also have an incentive to challenge the patent to avoid paying royalties to the patent holder.").

[65] *See generally* Taylor, *supra* note 22, at 235-41.

circumstances.[66] The other circuits have tended to take a more hostile attitude toward these clauses.[67] However, most cases from these courts tend to be of an older vintage, and judicial attitude may have since evolved.[68] The courts seem to distinguish between no challenge clauses based on the agreement they are embodied in. The exposition below will follow this practice.

2. No Challenge Clauses in Licensing Agreements

Two years after *Lear*, in *Massillon-Cleveland-Akron Sign Co. v. Golden State Advertising Co.*, the Ninth Circuit confronted a case involving an explicit no challenge clause in a settlement agreement.[69] In determining the validity of the clause, the Ninth Circuit made extensive reference to *Lear*. The court reiterated that the Supreme Court had struck the balance between state contract law and federal patent law, decisively in favor of promoting the federal patent policy of allowing the free flow of ideas that are not patented.[70] The Ninth Circuit was cognizant of the difference between the doctrine of licensee estoppel at issue in *Lear* and the no challenge clause at issue.[71] However, to the Ninth Circuit, this difference was immaterial. The court declared that "[t]he parties' contract, however, is no more controlling on this issue than is the State's doctrine of estoppel, which is also rooted in contract principles,"[72] and that the no challenge clause "is in just as direct conflict with the 'strong federal policy' referred to repeatedly in Lear, as was the estoppel doctrine and the specific contractual provision struck down in that decision."[73] Moreover, in dicta, the Ninth Circuit declared that for the purpose of the enforceability of no challenge clauses, there is

---

[66] *E.g.*, *Baseload Energy, Inc. v. Roberts*, 619 F.3d 1357 (Fed. Cir. 2010); *Flex-Foot v. CRP*, 238 F.3d 1362 (Fed. Cir. 2001).

[67] Massillon-Cleveland-Akron Sign Co. v. Golden State Advertising Co., 444 F.2d 425 (9th Cir. 1971); *Bendix*, 421 F.2d 809; Warner-Jenkinson Co. v. Allied Chemical Corp., 567 F.2d 184 (1977).

[68] *Massillon-Cleveland-Akron Sign*, 444 F.2d 425; *Bendix*, 421 F.2d 809; *Warner-Jenkinson*, 567 F.2d 184.

[69] *Massillon-Cleveland-Akron Sign*, 444 F.2d 425. The two parties to the case had been involved in a patent infringement dispute, which they settled in an agreement in 1962. In the agreement, the alleged infringer acknowledged the validity of the patent and that its action had infringed the patent. It further agreed not to challenge, directly or indirectly, the validity of the patent and not to infringe the patent again in the future. The agreement did not concern any licensing activity between the patentee and the alleged infringer.

[70] *Id*. at 425.

[71] *Id.* at 426.

[72] *Id.* at 427.

[73] *Id.*

no difference between a licensing agreement and a settlement agreement.[74] The court correctly recognized that a licensing agreement can be reached under the threat of a charge of infringement.[75] It also observed that such a distinction would be "less then [sic] workable,"[76] and would open the door to easy circumvention because "it would be just as easy to couch licensing arrangements in the form of settlement agreements."[77]

In *Bendix Corp v. Balax Inc.*, the Seventh Circuit struck down a no challenge clause in a licensing agreement that prohibited the licensees from challenging the validity of the patent even after the agreement had been terminated or lapsed.[78] The infringement defendants in that case alleged that the patentee used the no challenge clauses in the licensing agreements to "blanket" the market.[79] Citing *Lear* extensively, including the passage excerpted above, the court concluded that "the right to estop licensees from challenging a patent is not part of the 'limited protection' afforded by the patent monopoly."[80] More relevant to the purposes of this article, the court noted that the arrangement at issue should be struck down because "it creates a danger of unwarranted monopolization."[81] This danger was compounded by the fact that the obligation not to challenge extended beyond the duration of the licensing agreements.

More recently, in *Rates Technology v. Speakeasy, Inc.*,[82] the Second Circuit struck down a no challenge clause contained in a pre-litigation settlement agreement.[83] The court noted that what it was asked to do was "to balance the policy concerns of patent articulated in *Lear* against countervailing policy concerns that favor requiring parties to adhere to the terms of agreements resolving their

---

[74] *Id.*
[75] *Id.*
[76] *Id.*
[77] *Id.*
[78] Bendix Corp. v. Balax Inc., 421 F.2d 809 (7th Cir. 1970).
[79] *Id.* at 820.
[80] *Id.*
[81] *Id.*
[82] Rates Tech., Inc. v. Speakeasy, Inc., 685 F.3d 163 (2d Cir. 2012).
[83] The patentee plaintiff discovered an alleged infringement by the defendants and entered into a settlement agreement styled as a "Covenant Not to Sue" in 2007. After the entry of the agreement and a series of corporate transactions, the plaintiff patentee discovered continual infringement by the defendants and brought suit. In response one of the defendants sought a declaratory judgment action declaring that the plaintiff's patents were invalid. In a suit that eventually led to the appeal to the Second Circuit, the plaintiff alleged breach of contract by the defendants for violating the no challenge clause. *See generally id.*

legal disputes."[84] Citing *Massillon-Cleveland-Akron Sign Co.* with approval, the court observed that "allowing such no-challenges whenever a license agreement is cast as a 'settlement' could 'close the doors of the courts to a large group of parties who ha[ve] sufficient interest in the patent to challenge its validity,' [internal citation omitted] and thereby render Lear's prohibition of licensee estoppel—a prohibition that the Supreme Court held was required by strong public policy considerations—a dead letter."[85]

Importantly, while acknowledging that "the important policy interests favoring the settlement of litigation *may* support a different rule with respect to no-challenge clauses in settlements entered into after the initiation of litigation,"[86] the court held that "enforcing no-challenge clauses in pre-litigation settlements would significantly undermine the 'public interest in discovering invalid patents.'"[87] Despite the court's reference to the initiation of litigation, the court pronounced that the crucial watershed between enforceability and unenforceability is the conduct of discovery. No challenge clauses contained in a settlement agreement entered into after discovery would be enforceable, while those in a settlement agreement entered into prior to discovery would be void. To the court, discovery serves two important purposes:

> First, it suggests that the alleged infringer has had a full opportunity to assess the validity of the patent, and is therefore making an informed decision to abandon her challenge to its validity. Second, the fact that parties have conducted discovery is evidence that they had a genuine dispute over the patent's validity, and that the patent owner is not seeking to prevent its monopoly from being challenged by characterizing ordinary licensing agreements as settlement agreements.[88]

Because, as mentioned earlier, it is often impossible to draw the line between a pre-litigation settlement agreement and a licensing agreement,[89] the

---

[84] *Id.* at 171.
[85] *Id.*
[86] *Id.* at 172.
[87] *Id.*
[88] *Id.*
[89] While the timing between a pre-litigation settlement agreement and a licensing agreement may be difficult to distinguish, one may argue that the two differ by the presence or absence of a licensing arrangement. There need not be a licensing arrangement in a settlement agreement; the alleged infringer may merely agree to cease infringing activities. Meanwhile, a licensing agreement by definition must contain a licensing arrangement. This attempt at differentiation

court's conclusion effectively means that no challenge clauses in licensing agreements are unenforceable.

The position on no challenge clauses in licensing agreements would have been quite clear but for the Federal Circuit's decision in *Baseload Energy, Inc. v. Roberts*.[90] The court stated in dicta that "[i]n the context of settlement agreements, as with consent decrees, clear and unambiguous language barring the right to challenge patent validity in future infringement actions is sufficient, even if invalidity claims had not been previously at issue and had not been actually litigated."[91] The court made this statement while trying to distinguish the facts of the instant case from a prior case, *Flex-Foot v. CRP*.[92] In *Flex-Foot*, the alleged infringer had challenged patent validity, had had an opportunity to conduct discovery regarding validity, and had agreed voluntarily to dismiss the suit with prejudice in a settlement agreement containing a clear and unambiguous no challenge clause.[93] In *Baseload Energy*, the Federal Circuit argued that the exact factual circumstances need not be replicated for a no challenge clause to be upheld.[94] The focus seems to have shifted from the existence of prior litigation and prior opportunity to conduct discovery, which would have aligned the Federal Circuit with the Second Circuit, to the existence of clear and unambiguous language barring future validity challenges. This opens the possibility that the Federal Circuit would uphold a clear and unambiguous no challenge clause contained in a licensing agreement in the absence of any pending or ongoing litigation.

3. No Challenge Clauses in Settlement Agreements

The appellate courts have gone in different directions in their treatment of no challenge clauses contained in a settlement agreement. There seems to be some consensus that the dividing line for enforceability is whether the settlement

---

would be highly problematic for two reasons. First, settling parties that desire to enter into a licensing arrangement can easily circumvent the rule by inserting the no challenge clause in a settlement agreement while entering into a separate licensing agreement. Second, as a matter of policy, once one repudiates the rationale of licensee estoppel, it is unclear why the presence or absence of a licensing arrangement should have any bearing on the enforceability of a no challenge clause. Therefore, a better argument is that for the purpose of enforceability of no challenge clauses, pre-litigation settlement agreements and licensing agreements are to be treated the same.

[90] Baseload Energy, Inc. v. Roberts, 619 F.3d 1357 (Fed. Cir. 2010).

[91] *Id*. at 1363.

[92] Flex-Foot v. CRP, 238 F.3d 1362 (Fed. Cir. 2001).

[93] *Id.* at 1363-64.

[94] *Baseload Energy*, 238 F.3d at 1363.

agreement was entered into before or after discovery, or expense of substantial judicial resources. However, a number of cases deviate from this consensus. There are cases that held, or at least proclaimed, that no challenge clauses would be deemed unenforceable regardless of whether they are incorporated in a licensing agreement or a settlement agreement. There are also cases in which the court refused to enforce a no challenge clause, or at least something similar to it, contained in a settlement agreement entered into after discovery. Finally, there are also cases in which the court enforced a no challenge clause in a settlement agreement entered into prior to discovery.

A number of appellate decisions that have dealt with the enforceability of no challenge clauses in settlement agreements have upheld them so long as the settlement agreement was entered into after discovery. As mentioned, the Second Circuit in *Rates Technology* held that the dividing line for enforceability is discovery. In *Aro Corp. v. Allied Witan Co.*,[95] the Sixth Circuit enforced a no challenge clause in a settlement agreement entered into after discovery. Although the court did not explicitly designate discovery as the dividing line as the Second Circuit did in *Rates Technology*, it noted that *Lear* "cannot be interpreted so broadly as to condone a kind of gamesmanship, wherein an alleged infringer, after employing the judicial system for months of discovery, negotiation and sparring, abandons its challenge to validity, executes a license in settlement, and then repudiates the license and seeks to start the fight all over again in the courts."[96] The Federal Circuit has also consistently upheld no challenge clauses in settlement agreements that were entered into after discovery. *Hemstreet v. Spiegel, Inc.* did not concern an explicit no challenge clause.[97] It instead involved a provision that required the licensee to continue to pay royalty even after the patent had been otherwise invalidated, which in monetary terms functioned similarly as a no challenge clause. The Federal Circuit upheld the provision on the grounds of furthering settlement of lawsuits, despite the fact that the patent had been found unenforceable in a separate proceeding.[98] As noted earlier, in *Flex-Foot, Inc. v. CRP, Inc.*, the Federal Circuit upheld a no challenge clause in a settlement agreement entered into after discovery between two parties to an existing license.

---

[95] Aro Corp. v. Allied Witan Co., 531 F.2d 1368 (6th Cir. 1976).

[96] *Id.* at 1373. The Court did not emphasize the fact that discovery gave the settling parties sufficient information to make an informed decision. Instead, the Court believed that defendant had taken up so much judicial resources that it should not be given a second chance. *Id.*

[97] Hemstreet v. Spiegel, Inc., 851 F.2d 348 (Fed. Cir. 1988).

[98] *Id.* at 350.

The three cases that do not conform to this rough consensus were Massillon-Cleveland-Akron Sign Co. v. Golden State Advertising Co., Warner-Jenkinson Co. v. Allied Chemical Corp.,[99] and Baseload Energy, Inc. v. Roberts. However, it is possible to reconcile the first two cases with the general rule that discovery is the dividing line for enforceability. In Massillon-Cleveland-Akron Sign Co., the Ninth Circuit did indicate in dicta that no challenge clauses would be deemed unenforceable regardless of whether they are incorporated in a licensing agreement or a settlement agreement.[100] It is nonetheless important to note that the no challenge clause at issue in the case, which the court refused to enforce, was contained in a settlement agreement entered into prior to the commencement of litigation.[101]

In *Warner-Jenkinson*, the Second Circuit struck down a clause that prohibited a licensee from terminating the license for two years on the grounds that the licensee should be able to terminate the license if she successfully challenges the patent's validity.[102] While the clause was contained in an agreement reached by the parties after discovery in a prior litigation, the Court nonetheless refused to uphold it. However, this does not mean that the Court's holding is inconsistent with the general rule. In fact, the Court noted that if the agreement had contained an explicit no challenge clause, the Court may have felt compelled to give effect to it.[103] The Court merely observed that the *Lear* decision cautions against reading an explicit no challenge clause into an ambiguous clause such as the one at issue in the case.[104] Therefore, one may perhaps treat this case as not being applicable to explicit no challenge clauses at all.

Perhaps the one true anomaly among the three cases is *Baseload Energy*. In this case, the declaratory judgment defendant sought to enforce a claim release clause, under which the plaintiff has relinquished all present and future claims against the defendant, against the plaintiff.[105] The Federal Circuit ruled against the

---

[99] Warner-Jenkinson Co. v. Allied Chemical Corp., 567 F.2d 184 (2d Cir. 1977).

[100] Massillon-Cleveland-Akron Sign Co. v. Golden State Advertising Co., 444 F.2d 425, 427 (9th Cir. 1971).

[101] *Id.*

[102] *Warner-Jenkinson*, 567 F.2d at 188.

[103] *Id.*

[104] *Id.*

[105] Baseload Energy, Inc. v. Roberts, 619 F.3d 1357 (Fed. Cir. 2010).The parties had entered into a joint venture to develop some wind energy projects. The parties had reached an oral agreement concerning the terms of operation, but the joint venture broke down and one of the parties brought suit claiming breach of contract, fraud, and promissory estoppel. The parties settled the suit with an agreement which stipulated that both parties would release all claims

defendant, stating that the claim release clause did not specifically refer to invalidity issues and therefore could not be used to bar validity challenges.[106] However, in responding to the plaintiff's argument that the claim release clause should not bar its declaratory judgment action because the settlement agreement was not entered into after discovery and extensive court proceeding, the Court asserted that the absence of prior dispute or litigation as to invalidity is not dispositive of the enforceability issue.[107] If there was no prior dispute concerning invalidity, there clearly would have been no discovery on the issue. The Court implicitly noted that prior discovery on patent validity is not determinative of enforceability of no challenge clauses in settlement agreements.[108]

4. No Challenge Clauses in Consent Decrees

Given that this article will not focus on no challenge clauses in consent decrees, the discussion here will be brief. A few Federal Circuit cases can be interpreted as holding that a consent decree, which stipulates patent validity, bars future validity challenges absent express reservation of the right to launch such challenges. A majority of the appellate courts, however, have held that a consent decree (or a settlement agreement accompanied by a dismissal with prejudice) that stipulates patent validity and infringement precludes future validity challenges.[109]

---

against each other arising from any aspect of the venture. Their relationship broke down again, and one of the parties brought a declaratory judgment action, claiming that the patent that was to form the basis of the venture was invalid.

[106] *Id.* at 1363 (holding that clause did not contain clear and unambiguous language barring future validity challenges).

[107] *Id.*

[108] *Id.* ("In the context of settlement agreements, as with consent decrees, clear and unambiguous language barring the right to challenge patent validity in future infringement actions is sufficient, even if invalidity claims had not been previously at issue and had not been actually litigated.")

[109] Wallace Clark & Co. v. Acheson Indus., 532 F.2d 846, 849 (2d Cir. 1976) ("We conclude that the interests of litigants and the public in general will be best served by according res judicata effect to consent decrees adjudicating a patent's infringement as well as its validity."); American Equipment Corp. v. Wikomi Manufacturing Co., 630 F.2d 544, 548 (7th Cir. 1980) (noting that enforcing no challenge clauses in consent decrees is "the most effective way to enforce the Lear policy of facilitating competitive access to ideas".); Schlegel Mfg. Co. v. USM Corp., 525 F.2d 775, 780 (6th Cir. 1975) (noting that "there is a significant difference between the effect of a consent decree and the doctrine of licensee estoppel" in upholding no challenge clauses in consent decrees); Kraly v. National Distillers & Chem. Corp., 502 F.2d 1366, 1370 (7th Cir. 1974) (noting that federal patent policy "must occupy a subsidiary position to the fundamental policy favoring the expedient and orderly settlement of disputes and the fostering of judicial economy").

In other words, a no challenge clause stipulated in such a consent decree would be enforceable.

5. Termination-Upon-Challenge Clauses

The courts' attitudes toward termination-upon-challenge clauses is similar to that toward general no challenge clauses, in that the Federal Circuit holds a more lenient position than the other circuits. In *Crane Co. v. Aeroquip Corp.*, the Seventh Circuit held termination-upon-challenge clauses to be unenforceable on the grounds that under *Lear*, "[d]efendant was within its rights to test validity after entering into the consent judgment of validity."[110] However, in *C.R. Bard, Inc. v. Schwartz*, the Federal Circuit implicitly held that a licensor can terminate the licensing agreement when a licensee sues to declare the patent invalid and ceases to pay royalty.[111] Despite the slightly complex facts in *Schwartz*, commentators have argued that in so ruling, "the Federal Circuit effectively held that license provisions which give licensors the right to terminate licenses are enforceable when licensees bring validity challenges and cease making royalty payments."[112] Taylor argues that Federal Circuit case law such as *Cordis Corp. v. Medtronic, Inc.* lends further support to the notion that termination-upon-challenge clauses should be enforceable.[113] The gist of the Federal Circuit's approach is that while *Lear* requires the courts to allow the licensee to challenge the validity of the patent, the licensee should not be spared of the consequences of a validity challenge.[114] The licensee should not be allowed to launch a validity challenge while continually enjoying the benefit of the licensing agreement. The implication would be that the patentee should be allowed to terminate the licensing agreement, at least when the licensee also ceases to pay royalty.[115] If termination-upon-challenge clauses are more likely to be upheld by the courts and are equally effective in deterring validity challenges, one may see them incorporated in licensing and settlement agreements more often, and they may end up featuring more prominently in antitrust cases.

---

[110] Crane Co. v. Aeroquip Corp., 504 F.2d 1086, 1092 (7th Cir. 1974).

[111] C.R. Bard, Inc. v. Schwartz, 716 F.2d 874 (Fed. Cir. 1983).

[112] Taylor, *supra* note 22, at 249.

[113] *Id.* at 250-51.

[114] Miller & Gal, *supra* note 6, at 133.

[115] It would be interesting to see is whether a patentee can contractually stipulate a right to terminate the licensing agreement when a licensee launches a validity challenge while still paying royalty. That would give the patentee an unqualified right to terminate the licensing agreement, as opposed to under *Schwartz*, where the licensee could still try to maintain the licensing agreement by continuing to pay royalty.

B. The European Union

Unlike the U.S., the EU has dealt with no challenge clauses under competition law. Given the fact that patent law is still largely national law in the EU, relegating the treatment of no challenge clauses to patent law, like it has been done in the U.S., could result in a variety of approaches. A review of the EU approach to no challenge clauses entails an examination of both the case law of the European courts and the European Commission's TTBER. On the whole, it is fair to say that both the European courts and the European Commission have taken a fairly hardline approach toward no challenge clauses, even though their approaches, especially that of the Commission, have evolved over time.

1. The European Courts

The European Court of Justice (ECJ), now renamed the Court of Justice of the European Union (CJEU), encountered no challenge clauses in *Windsurfing International v. Commission*.[116] In that case, the patentee imposed an express no challenge clause on its licensees, which the European Commission challenged as being incompatible with Article 101(1) (then Article 85(1)) of the TFEU. The ECJ condemned the clause in summary fashion, without an examination of competitive effects. Although the court stopped short of ruling that no challenge clauses restrict competition by object, it did hold that no challenge clauses infringe Article 101(1) because of the overriding public interest in removing invalid patents, without any regard to possible competitive effects.[117] The court further concluded that the clause did not benefit from the exemption under Article 101(3).[118] Using U.S. antitrust parlance, commentators have remarked that the ECJ in *Windsurfing* condemned the no challenge clause as illegal per se.[119]

---

[116] Case C-193/83 Windsurfing Int'l, Inc. v. Commission, 1986 E.C.R. 611.

[117] *Id.* at 663.

[118] The result in this case is all the more remarkable because the validity of the patent at issue had been closely examined by the German courts and no licensees had brought a validity challenge after the clause had been dropped in response to Commission action. Orstavik, *supra* note 7, at 103. In other words, the ECJ condemned a no challenge clause when the underlying patent was in all likelihood valid and when it was clear that none of the licensees had been prevented by the clause from bringing challenges as none had the incentive to do so. *Windsurfing Int'l*, E.C.R. 611 at 664. This is practically tantamount to saying that no challenge clauses are illegal even though their incorporation into a licensing agreement has no impact on eventual patent validity or licensee incentive to challenge.

[119] Orstavik, *supra* note 7, at 103; P. Sean Morris, *Patent Licensing and No Challenge Clauses: A Thin Line Between Article 81 EC Treaty and the New Technology Transfer Block Exemption Regulation*, 3 INTELL. PROP. Q. 217, 221-22 (2009).

The ECJ modified its position on no challenge clauses in a subsequent case, *Bayer v. Süllhöfer*.[120] The court began its discussion by rejecting the Commission's two arguments. First, the court rejected the argument that no challenge clauses are, in principle, to be considered a restriction of competition under Article 101(1). Second, the court disagreed with the Commission's argument that these clauses can be compatible with Article 101(1) if they are incorporated in a settlement agreement and some further conditions are met. The ECJ held that as far as no challenge clauses are concerned, it makes no difference whether they are in a licensing agreement or a settlement agreement.[121] Instead, the court held that one must take into account "the legal and economic context" in determining the legality of these clauses.[122] The court proceeded to enumerate two circumstances in which no challenge clauses would be permissible: (1) when the license that contains a no challenge clause is free, which means that the licensee does not suffer from the competitive disadvantage of royalty payment, and (2) "when the licence relates to a technically outdated process which the licensee undertaking did not use."[123]

In the more recent *Huawei Technologies v. ZTE Corp.* case,[124] the CJEU had another opportunity to discuss the importance of the right of a licensee to challenge the validity of the licensed patent. In this case, Advocate General Wathelet stated in his opinion that:

> [I]t is in the public interest for an alleged infringer to have the opportunity, after concluding a licensing agreement, to challenge the

[120] Case 65/86, Bayer AG v. Süllhöfer, 1988 E.C.R. 5249.

[121] *Id.* at 5286. However, the ECJ reserved the question of whether no challenge clauses are illegal if they are incorporated in a consent decree.

[122] *Id.*

[123] *Id.* The ECJ's articulation of the first exception reveals a misunderstanding of the competitive harm of no challenge clauses and represents a peculiar and unexplained departure from the rationale articulated in *Windsurfing* for condemning no challenge clauses. The ECJ seemed to believe that no challenge clauses harm competition because they prevent licensees from extricating themselves from a royalty payment obligation if the patent turns out to be invalid. They deprive the licensees of an opportunity to challenge the patent. However, that is not the reason why no challenge clauses harm competition. Instead, they harm competition because they allow the patentee artificially to maintain the market power that it may have obtained from an invalid patent, when the patentee should be entitled to no such market power. In *Windsurfing*, the rationale for invalidating no challenge clauses is the public interest in clearing the market of invalid patents. Here, the rationale seems to have shifted to protecting licensees from unjustified royalty payments.

[124] Case C-170/13, Huawei Technologies Co. v. ZTE Corp., 2015 Curia ECLI:EU:C:2015:477 (July 16, 2015).

> validity of an SEP (as ZTE did). As the Commission has pointed out, the wrongful issue of a patent may constitute an obstacle to the legitimate pursuit of an economic activity. Moreover, if undertakings supplying standard-compliant products and services cannot call into question the validity of a patent declared to be essential to that standard, it could prove effectively impossible to verify the validity of that patent because other undertakings would have no interest in bringing proceedings in that regard.[125]

This case has the added dimension of involving standard-essential patents ("SEPs"), which have more serious competitive implications because they tend to possess substantial market power. Echoing the U.S. Supreme Court's observation in *Lear*, Advocate General Wathelet speculated that licensees may be the only party with the incentive to challenge the validity of a patent.[126] The CJEU agreed with Advocate General Wathelet, and stated in its judgment that:

> [H]aving regard, first, to the fact that a standardisation body such as that which developed the standard at issue in the main proceedings does not check whether patents are valid or essential to the standard in which they are included during the standardisation procedure, and, secondly, to the right to effective judicial protection guaranteed by Article 47 of the Charter, an alleged infringer cannot be criticised either for challenging, in parallel to the negotiations relating to the grant of licences, the validity of those patents and/or the essential nature of those patents to the standard in which they are included and/or their actual use, or for reserving the right to do so in the future.[127]

Particularly noteworthy is the Court's observation that standard setting organizations ("SSOs") do not necessarily check the validity or essentiality of the patents seeking to be included in the standards.[128] In fact, most SSOs do not check the validity of the included patents.[129] Given the fact that standardization would

---

[125] Case C-170/13, Huawei Technologies Co. v. ZTE Corp., 2014 Curia ECLI:EU:C:2014:2391 *95 (Nov. 20, 2014).

[126] *Id.* at *64.

[127] *Id.*

[128] Case C-170/13, Huawei Technologies Co. v. ZTE Corp., 2015 Curia ECLI:EU:C:2015:477 *69 (July 16, 2015).

[129] NATIONAL RESEARCH COUNCIL (U.S.) COMM. ON INTELL. PROP. MGM'T IN STANDARD-SETTING PROCESSES, PATENT CHALLENGES FOR STANDARD-SETTING IN THE GLOBAL ECONOMY: LESSONS FROM INFORMATION AND COMMUNICATIONS TECHNOLOGY § 7.2 (2013).

give patents a great deal of market power, the harm of allowing an invalid SEP to persist is much greater than for a non-SEP.

2. The European Commission

The European Commission has had significant influence over the licensing practices of European patent holders. Patentees usually try to steer clear of what the Commission deems to be impermissible licensing practices to take advantage of the legal certainty provided by the TTBER.[130] The Commission's view of no challenge clauses has evolved over time. In 2014, the European Commission issued the most recent set of TTBER.[131] In these regulations, the Commission revised its position on no challenge clauses, which continue to be an excluded restriction, meaning they will not automatically benefit from the block exemption and their compatibility with the Treaty will have to be individually assessed.[132] However, termination-upon-challenge clauses are now classified as an excluded restriction as well, which previously were not under the 2004 TTBER, except when incorporated in an exclusive license and the market share thresholds provided in Article 3 of the TTBER are met.[133] In the accompanying guidelines, the Commission asserted that no challenge clauses are likely to fall within Article 101(1) when the licensed technology is valuable, and therefore creates a competitive advantage for the licensees.[134] In such a case, a no challenge clause is unlikely to meet the conditions for Article 101(3).[135] This means that it would be outright illegal, which is reminiscent of the ECJ's position in *Windsurfing*. Finally, the Commission incorporated the two exceptions provided by the ECJ in *Bayer v. Süllhöfer*.[136]

In the 2014 TTBER, the Commission made two major changes to its position on no challenge clauses. First, it took a slightly more cautious approach to no challenge clauses in settlement agreements.[137] After repeating its previous

---

[130] Lawrence, *supra* note 5, at 811.

[131] *Id*. at 810.

[132] 2014 TTBER, *supra* note 51, at art. 5(1)(b).

[133] *Id*. Article 3 provides that the exemptions provided in the TTBER apply when the aggregate market share of parties to an agreement which share a horizontal relationship is no more than 20%, and that of parties which share a vertical relationship is no more than 30%.

[134] Guidelines on the Application of Article 101 of the Treaty on the Functioning of the European Union to Technology Transfer Agreements, 2015 O.J. C89/3, art. 114 [hereinafter "Guidelines on the 2014 TTBER"].

[135] *Id*. at 134.

[136] *Id.*

[137] It is important to note that the TTBER only applies to licensing agreements. Therefore, the TTBER would only apply to a settlement agreement that includes licensing provisions.

position that no challenge clauses in settlement agreements generally fall outside Article 101(1), it proceeded to caution that these clauses nonetheless could be anticompetitive under specific circumstances.[138] Second, the Commission included termination-upon-challenge clauses, except in exclusive licenses, as an excluded restriction. The Commission explained this change of position by saying that: "[s]uch a termination right can have the same effect as a non-challenge clause, in particular where switching away from the licensor's technology would result in a significant loss to the licensee … or where the licensor's technology is a necessary input for the licensee's production."[139] The key factor to consider is whether the loss of profit would act as a sufficient deterrent to challenges, which, according to the Commission, will need to be assessed on a case-by-case basis.[140]

C. China

1. Regulations and Draft Guidelines

Of the three jurisdictions surveyed in this article, China appears to take the strictest approach to no challenge clauses, as evidenced in the decisional practices of Chinese enforcement authorities. This may reflect a strategic concern that most of China remains a net importer of foreign technologies.[141] A more pro-licensee approach would stand to benefit Chinese companies. The three Chinese enforcement authorities, the NDRC, the SAIC, and the Ministry of Commerce,[142]

---

[138] Guidelines on the 2014 TTBER, *supra* note 134, at 243. The Commission raised the example of when the intellectual property right was granted following the submission of incorrect or misleading information and when the technology rights are a necessary input in the licensee's production.

[139] *Id*. at art. 136.

[140] *Id*. As examples of such a possible case, the Commission mentioned situations when the patents being licensed are standard-essential and when the licensed technology has a very significant market position.

The Commission further justified the special treatment for termination-upon-challenge clauses in an exclusive licensing agreement on the grounds that "the incentives for innovation and for licensing out could be undermined if, for example, the licensor were to be locked into an agreement with an exclusive licensee which no longer makes significant efforts to develop, produce and market the product (to be) produced with the licensed technology rights." *Id*. at 139. An alternative to allowing termination-upon-challenge clauses is to allow the patentee to cancel the exclusivity provision, which would solve the problem of a patentee being locked in by an uncooperative licensee.

[141] Hejing Chen & John Whalley, *China's Post-1978 Growth Process and Earlier Growth Processes of Europe, US, Japan, and Korea, in* WORLD SCIENTIFIC REFERENCE ON ASIA AND THE WORLD ECONOMY VOL. 2, 44 (Manmohan Agarwal & John Whalley eds).

[142] MOFCOM is responsible for merger reviews and therefore probably has less to do with no challenge clauses. The NDRC and the SAIC share responsibility in conduct enforcement.

are reportedly drafting the IP-Competition Guidelines under the auspices of the Anti-Monopoly Commission, which is an advisory body in the State Council overseeing and coordinating enforcement activity by the three authorities.[143] This article will make reference to the approach to no challenge clauses in the consultative drafts released by the NDRC and the SAIC. It will also refer to the *Regulation on the Prohibition of Conduct Eliminating or Restricting Competition by Abusing Intellectual Property Rights* released by the SAIC in April 2015 ("the "SAIC Regulation").[144] The SAIC Regulation also contains some discussion of no challenge clauses.[145] Lastly, and most importantly, in February 2015, the NDRC found that Qualcomm had abused its dominance through a variety of licensing practices, including imposing no challenge clauses on its licensees.[146] Qualcomm was fined RMB6 billion (approximately USD1 billion).[147] An examination of the decision will shed light on the prevailing Chinese approach to no challenge clauses.

Article 10 of the SAIC Regulation stipulates that a business operator with a dominant market position should not, without legitimate reasons, prohibit transaction counterparties from raising doubts about the validity of its intellectual property rights, thereby eliminating or restricting competition.[148] Although the article uses the phrase "raising doubts about", it probably refers to launching a validity challenge. Otherwise, the language is so impermissibly broad that the SAIC could not have reasonably contemplated that interpretation.

---

[143] The current approach is that each of the authorities will come up with its own draft and the Anti-Monopoly Commission will be responsible for combining them. As of the time of writing, the NDRC and SAIC have released consultative drafts, which are likely to be similar to their final drafts. The expectation is that the final combined product will be released in June.

[144] Provisions on Prohibiting the Abuse of Intellectual Property Rights to Exclude and Restrain Competition, *supra* note 53.

[145] *Id.* at art. 10.

[146] National Development and Reform Commission Administrative Penalty Decision [2015] No. 1, Feb. 9, 2015, http://jjs.ndrc.gov.cn/fjgld/201503/t20150302_666170.html (hereinafter "Qualcomm Decision").

[147] *Id.*

[148] *Id.* at art. 10. One notable feature about SAIC's approach to no challenge clauses is that instead of treating them as a potential restrictive agreement, the focus is on treating them as an abuse of dominance. The implicit recognition seems to be that no challenge clauses would not cause competitive harm unless they are imposed by a firm with substantial market power. Whether the amount of market power necessary to create competitive harm must reach the level of dominance is open to debate, but it will be clear from subsequent discussion that market power is necessary for no challenge clauses to inflict competitive harm.

A few distinctions are key to understanding Article 10. First, the provision refers to transactional counterparty and not licensee. Therefore, it probably has a broader reach than the EU TTBER, and could potentially cover buyers of products that incorporate the patented technology in addition to licensees. Second, the provision is worded in such a way that it seems to require a demonstration of a restriction of competition before a no challenge prohibition will be outlawed. If this is true, it appears that the SAIC has not adopted a per se approach to no challenge clauses, and may be more lenient with them than the European Commission. However, it is not clear how much importance should be attached to the language "elimination or restriction of competition". There have been cases in the past in which the enforcement authorities' guidelines indicated that the conduct at issue requires a showing of competitive effects, but the authorities did not make such a showing in their decisions. Third, the provision does provide for the possibility of justification by way of "legitimate reasons," even though it stops short of defining these reasons. This may be further evidence that the SAIC does not adopt a per se approach. Finally, the provision refers only to a prohibition of challenges by transactional counterparties. At least on a literal interpretation, it does not seem to cover provisions such as termination-upon-challenge clauses or higher royalty-upon-challenge clauses that stop short of outright prohibiting challenges, but merely create hurdles for them. The seventh consultative draft of the IP-Competition Guidelines issued by the SAIC by and large repeats the same language as the SAIC Regulation regarding no challenge clauses.[149]

The NDRC draft IP-Competition Guidelines ("NDRC Guidelines") provide more detail on no challenge clauses.[150] The first notable feature about these guidelines is that no challenge clauses are discussed under both the restrictive agreements section and the abuse of dominance section.[151] In the restrictive agreements section, the guidelines provide a relatively detailed discussion about these clauses. Article 2(1)(3) begins by acknowledging that no challenge clauses can serve the useful purposes of preventing excessive litigation and improving

[149] Guanyu Lanyong Zhishi Chanquan de Fanlongduan Zhifa Zhinan (Guojia Gongshang Zhongju Diqigao) (关于滥用知识产权的反垄断执法指南（国家工商总局第七稿）) [Guidelines on Anti-Monopoly Enforcement against Abuse of Intellectual Property Rights (SAIC Seventh draft) (released by State Admin. of Trade and Com., Feb 4, 2016) (on file with author).

[150] Guowuyuan Fanlongduan Weiyuanhui Guanyu Lanyong Zhishi Chanquan de Fanlongduan Zhinan (Zhengqiu Yijiangao) (国务院反垄断委员会 关于滥用知识产权的反垄断指南（征求意见稿) [State Council Anti-Monopoly Commission Anti-Monopoly Guidelines on Abuse of Intellectual Property Rights (consultative draft)] (released by Nat'l Dev. and Reform Comm'n, Dec. 31, 2015) (on file with author) (hereinafter "NRDC Guidelines").

[151] *Id.* at art. 2(1)(3) & art. 3(2)(4).

transactional efficiency.[152] The section then proceeds to assert that these clauses can also restrict competition, which is to be determined with reference to a number of factors, including: (1) whether the patentee imposes the no challenge clause on all licensees, (2) whether the underlying patent is being licensed for royalty and whether the patent may constitute entry barriers into the downstream market, (3) whether the underlying patent blocks the implementation of other competing patents, (4) whether the patentee obtained the patent by providing false or misleading information, and (5) whether the patentee compels the licensee to accept the no challenge clause through improper means.[153] In the abuse of dominance section, Article 3(2)(4) merely lists the prohibition of licensees from challenging the licensor's patent as a prohibited unreasonable licensing condition, without any explanation of the relevant factors to be considered.[154] It is not entirely clear what explains the different treatment of no challenge clauses in the two sections. It may mean that if these clauses are treated as an abuse of dominance, the analytical process is simpler and there is no need to resort to the factors listed in Article 2(1)(3). Alternatively, it may simply mean that the factors listed in Article 2(1)(3) are tacitly incorporated in Article 3(2)(4). The latter explanation seems to make more sense, as there is no reason why different analytical factors are considered when no challenge clauses are treated as a restrictive agreement as opposed to an abuse of dominance.

2. The NDRC Qualcomm Decision

Apart from the SAIC Regulation and these draft guidelines, there has been one enforcement action that concerns no challenge clauses. In the NDRC's decision against Qualcomm released in February 2015, one of the four claims raised by the NDRC is the imposition of unreasonable conditions on the sale of the baseband chips used in mobile communication terminals.[155] One of the unreasonable conditions is that Qualcomm will terminate the supply of chips if the licensee initiates litigation against it,[156] which the NDRC characterizes as a no challenge clause.[157] Because Qualcomm stopped short of outright prohibiting

[152] *Id*. at art. 2(1)(3).

[153] *Id.*

[154] Similar to the SAIC Regulation, Article 3(2)(4) of the draft NDRC Guidelines only refers to outright prohibition. Therefore, it is again not clear whether it encompasses challenge-penalty clauses. In contrast, Article 2(1)(3) refers to no challenge clauses in general. Although the NDRC Guidelines do not define no challenge clauses, the general reference means that it may encompass challenge-penalty clauses as well.

[155] Qualcomm Decision, *supra* note 146.

[156] *Id.*

[157] *Id.*

licensees from initiating litigation,[158] the provision was, at most, a challenge-penalty clause. However, the NDRC was convinced that cessation of supply was enough of a deterrent to Qualcomm's customers, the terminal manufacturers, and thus, practically functioned as an outright prohibition.[159] Qualcomm admitted to the imposition of no-challenge clauses in the licensing agreements, but argued that its conduct was justified. The NDRC did not detail what the justifications were, but dismissed them as insufficient. According to the NDRC, it is within the licensee's' right to challenge patent validity or institute litigation with respect to the licensing agreements.[160] Qualcomm's imposition of no-challenge clauses restricted, if not outright deprived, the licensees of this right. Moreover, the NDRC argued that competition was restricted when potential licensees that were unwilling to accept the no-challenge clauses were excluded from the market.[161]

The NDRC did not consider the competitive effects of no challenge clauses except by saying that licensees that are unwilling to accept the unreasonable licensing terms would be excluded from the market.[162] However, that would be tantamount to saying that any time Qualcomm turns away a potential licensee, there is restriction of competition. The NDRC's alternative argument that no challenge clauses infringe upon the licensee's right to challenge patent validity effectively means that these clauses are illegal on their face. It would therefore seem that the NDRC's approach to no challenge clauses is stricter than that manifested in the SAIC Regulation and possibly in line with the approach taken in the 2014 TTBER.

---

[158] *Id.*

[159] *Id.*

[160] *Id.* (noting that initiating litigation over licensing agreement is licensee's right).

[161] This characterization of restriction of competition is highly problematic. Restriction of competition cannot be established simply on the grounds that Qualcomm offered some licensing terms which were unacceptable to some licensees. The corollary of this argument would be that Qualcomm would be expected to accept whatever terms its licensees are willing to offer. This certainly cannot be the case. What stops a potential licensee from arguing that competition is restricted when it fails to obtain a license because the licensee is unwilling to pay the license fee Qualcomm demands? Whether there is restriction of competition crucially depends on whether the licensing term demanded by Qualcomm is itself anticompetitive, which is what needs to be established in the analysis. The NDRC merely states that the clause denies the licensees their right to challenge the validity of the licensing agreements. Such a right-based argument fails to consider the competitive effects of the clause. The NDRC's analysis is thus found wanting.

[162] Qualcomm Decision, *supra* note 146 (noting that those unwilling to accept Qualcomm's licensing terms were excluded from market).

III
CONSUMER HARM OF NO CHALLENGE CLAUSES

In light of the differing approaches to no challenge clauses taken in these three jurisdictions, it is worth considering what the correct approach to these clauses should be. While the U.S. has largely regulated no challenge clauses under patent law, the EU and China have expressly subjected no challenge clauses to competition law. Nothing in U.S. antitrust law says that no challenge clauses are exempted from antitrust scrutiny. Nonetheless, as far as this author is aware, U.S. courts have not had the opportunity to rule on their legality under antitrust law. No challenge clauses also received no mention in the 1995 DOJ-FTC IP-Antitrust Guidelines.[163] There was brief mention of these clauses in the report issued by the DOJ and the FTC on IP-antitrust issues in 2007,[164] which states that "[i]nvalid patents impair competition, and as a matter of patent policy, challenges to their validity are encouraged."[165] It is noteworthy that the report cited to *Lear, Inc. v. Adkins* and *MedImmune Inc. v. Genentech, Inc.*, and not an antitrust case, as support for this statement. The report further refers to the Solicitor General's brief in *MedImmune* for the observation that "public policy strongly favors ridding the economy of invalid patents, which impede efficient licensing, hinder competition, and undermine incentives for innovation."[166] There is no allusion to consumer harm resulting from these clauses. Therefore, no challenge clauses are viewed through the lens of patent policy as opposed to antitrust policy. Meanwhile, the seemingly strict approach to no challenge clauses under EU and Chinese competition law would suggest that these two jurisdictions believe that these clauses can inflict considerable consumer harm that warrants the scrutiny of competition law.

Therefore, the first question to consider is whether no challenge clauses inflict harm on consumers. A short answer is that they do. Commentators have noted that no challenge clauses can create consumer harm under certain circumstances. Morris notes that "[t]he competitive harm associated with a no-challenge clause involves the risk that invalid intellectual property rights give their holders market power that is not justified by the policies underlying those rights.

---

[163] US DEP'T OF JUSTICE & FEDERAL TRADE COMM'N, ANTITRUST GUIDELINES FOR THE LICENSING OF INTELLECTUAL PROPERTY (1995), http://www.justice.gov/sites/default/files/atr/legacy/2006/04/27/0558.pdf.

[164] US DEP'T OF JUSTICE & FEDERAL TRADE COMM'N, ANTITRUST ENFORCEMENT AND INTELLECTUAL PROPERTY RIGHTS: PROMOTING INNOVATION AND COMPETITION (2007), http://www.justice.gov/sites/default/files/atr/legacy/2007/07/11/222655.pdf.

[165] *Id*. at 90.

[166] *Id*. at 90-91.

Such concentration of market power may lead to higher prices or lower output."[167] Likewise, Orstavik observes that under a no challenge clause, a "licensee may be obliged to pay royalties when none are justified, or the agreement may contain other restrictions that continue to apply even if the original right is invalid, thus restricting competition. Because of the obligation to pay royalties, the no-challenge clause may contribute to an artificially high price level."[168]

A. A Comparison with Reverse Payments

Reverse payments, in which the infringement plaintiff agrees to pay the defendant compensation, usually a very large sum of money, to settle the infringement suit, share important similarities with no challenge clauses. They both arise in the context of patent settlements (in no challenge clauses, also in licensing agreements), both entail the infringement defendant acknowledging the validity of the contested patent, and the legality of both practices hinge on patent validity. There are admittedly crucial differences between them, one of which being that, while reverse payments entail a large transfer from the infringement plaintiff to the infringement defendant, that need not be the case with no challenge clauses.[169] Therefore, if reverse payments have been roundly perceived to have serious anticompetitive potential—in fact, notable commentators have urged that they be held presumptively illegal[170]—no challenge clauses should at least deserve some antitrust scrutiny.

While acknowledging the similarities between reverse payments in pharmaceutical settlements[171] and licensing agreements, which may contain no

---

[167] Morris, *supra* note 119, at 220.

[168] Orstavik, *supra* note 7, at 102.

[169] Einer Elhauge & Alex Krueger, *Solving the Patent Settlement Puzzle*, 91 TEXAS L. REV. 283, 284 (2012)

[170] Herbert Hovenkamp, Mark Janis & Mark A. Lemley, *Anticompetitive Settlement of Intellectual Property Disputes*, 87 MINN. L. REV. 1719 (2003); Elhauge & Krueger, *supra* note 169; Michael A. Carrier, *Unsettling Drug Patent Settlements: A Framework for Presumptive Illegality*, 108 MICH. L. REV. 3 (2009).

[171] Reverse settlements, also known as pay-for-delay agreements, are settlement agreements, usually arising in the pharmaceutical industry, in which the patentee pays the potential infringer to settle the lawsuit. Because of the unusual direction of payment—one would usually expect the potential infringer to pay the patentee in a normal settlement agreements, reverse payments have received considerable attention from the courts and commentators over the years. The Supreme Court finally decided in *FTC v. Actavis* that reverse payments are subject to antitrust law and are to be analyzed under the Rule of Reason. Fed. Trade Comm'n v. Actavis, 570 U.S. 756 (2013). For academic commentary on the appropriate treatment of reverse payments under antitrust law, see Carl Shapiro, *Antitrust Limits to Patent Settlements*, 34(2) RAND J. ECON. 391 (2003);

challenge clauses, Miller and Gal justify their disparate treatment under antitrust law. They highlight a number of major differences between reverse payments and no challenge clauses that justify the hands-off approach of U.S. antitrust law to the latter. First, licensing agreements "are ongoing, they may further social welfare … , and they are generally based on the assumption that the patent is valid, at least when the contract is signed."[172] Second, their effect on competition is different in that "[l]icensees already operate in the market, albeit with restrictions contained in their licences, so the anticompetitive harm stems from restricting the entry of third parties into the market. By contrast, in pay-for-delay agreements [reverse payments] the harm includes the prevention of entry of the potential patent challenger."[173]

There are a number of problems with Miller and Gal's arguments. First, it is not clear how distinct reverse payment agreements are from licensing agreements containing no challenge clauses. Both types of agreements are ongoing, and their legal obligations persist for the duration of the agreements. The difference is that reverse payment agreements involve ongoing inaction, where the obligation is to abstain from the market, whereas licensing agreements with no challenge clauses involve ongoing action. Here, the ongoing activity is the commercialization of the patented technology through licensing and the ongoing obligation is to refrain from mounting a validity challenge. Even if licensing agreements could be construed as more more ongoing in nature than reverse payment agreements, it is not clear what the relevance of that is to consumer harm, so long as the harm is continuous under both agreements.

Second, licensing agreements may further social welfare by encouraging the commercialization of technology, whereas the only conceivable social benefit of reverse payments is the minimization of litigation. However, the correct comparison with reverse payments are not licensing agreements, but no challenge clauses. No challenge clauses themselves do not promote the commercialization of technology (unless one argues that the patentee will not license the technology absent these clauses, which will be addressed subsequently). The only purpose they serve, like reverse payments, is the minimization of disputes over patent validity.

Murat C. Mungan, *Reverse Payments, Perverse Incentives*, 27 HARV. J. L. & TECH. 1 (2013); Daniel A. Crane, *Exit Payments in Settlement of Patent Infringement Lawsuits: Antitrust Rules and Economic Implications*, 54 FLA. L. REV. 747 (2002); Thomas F. Cotter, Commentary, *Refining the "Presumptive Illegality" Approach to Settlements of Patent Disputes Involving Reverse Payments: A Commentary on Hovenkamp, Janis & Lemley*, 87 MINN. L. REV. 1789 (2003).

[172] Miller & Gal, *supra* note 6, at 136, n. 76.

[173] *Id*.

Furthermore, even if we were to compare licensing agreements and reverse payments, licensing agreements are not immune to anticompetitive uses. The patentee may intentionally disguise reverse payments in the form of a reduced royalty by undercharging the licensee.[174] While this mechanism would most likely be less effective than a lump sum transfer from the patentee, it has the advantage of being more difficult to detect and police. In order to show that there is a reverse payment, the court would need to establish what the royalty would be without the disguised reverse payment, which would be very difficult.[175] In the aftermath of *Actavis*,[176] we are already witnessing reverse payment agreements that eschew lump sum transfers but instead resort to complicated licensing, co-marketing, or delayed entry arrangements.[177]

Third, while it may be true that licensing agreements are generally premised on the validity of the underlying patent, what matters is not whether the practice at issue is premised on patent validity, which only pertains to the subjective state of mind of the parties, but whether the legality of the practice turns on patent validity. That is the relevant issue as far as antitrust analysis is concerned. The legality of reverse payments would turn on the validity of the underlying patent. Reverse payments are only objectionable as a matter of antitrust law if the underlying patent is invalid.[178] If the patent was invalid, the patentee would be effectively splitting with the potential infringer the monopoly profit, which she does not deserve. The patentee and the potential infringer are both better off than if the potential infringer enters the market after invalidating the patent. Monopolist profit

---

[174] Crane, *supra* note 171, at 776.

[175] Damien Geradin, Douglas Ginsburg & Graham Safty, *Reverse Payment Patent Settlements in the European Union and the United States*, GEO. MASON U. LEGAL STUDIES RES. PAPER SERIES LS 15-22 8, http://www.law.gmu.edu/assets/files/publications/working_papers/LS1522.pdf.

[176] Fed. Trade Comm'n v. Actavis, 570 U.S. 756 (2013).

[177] William O. Kerr & Cleve B. Tyler, *Measuring Reverse Payments in the Wake of* Actavis, 28 ANTITRUST 29, 30 (2013). See also King Drug Co. of Florence, Inc. v. Smithkline Beecham Corp. d/b/a GlaxoSmithKline, et al., 791 F.3d 388 (3d Cir. 2015); *In re* Loestrin 24 Fe Antitrust Litig., 45 F.Supp.3d 180 (D.R.I. 2014). In King Drug, GSK, the branded manufacturer, and Teva, the generic manufacturer, reached a settlement agreement that permitted Teva to enter the market for lamotrigine chewables 37 months early and the market for lamotrigine tablets 6 months early. The former agreed to stay out of the authorized generic market for lamotrigine during Teva's exclusivity period from July 2008 to January 2009. In return, Teva dropped its challenge to GSK's Lamictal patents. In In re Loestrin, Watson, the generic manufacturer, agreed to halt its patent validity challenge and abstain from the market for a time in exchange for the branded manufacturer Warner Chilcott's promise that, when Watson finally did enter, Warner Chilcott would refrain from competing with its own authorized generic version for six months.

[178] Shapiro, *supra* note 171, at 397.

is always higher than the profit redounding to two competing duopolists.[179] Furthermore, invalidating the patent would expose the potential infringer to further entry by third parties. Meanwhile, if the patent were valid, the patentee would have no reason to pay the potential infringer anything, but would be free to split its monopoly profit as she sees fit. In this situation, the reverse payment may be irrational from the patentee's perspective, but certainly would not be illegal. The patentee is entitled to exclude a potential infringer by exercising its patent right anyway. Similarly, no challenge clauses are only problematic if the underlying patent is invalid. By protecting an invalid patent, no challenge clauses augment the likelihood that invalid patents persist in the market and cause supra-competitive prices. If, however, the underlying patent were valid, the only consequence of no challenge clauses would be to eliminate needless litigation that would result in affirmation of patent validity anyway. This would be a socially beneficial outcome.

Lastly, Miller and Gal argue that reverse payment agreements exclude both the potential infringer and third parties, whereas licensing agreements with a no challenge clause only restrict the entry of third parties into the market.[180] Whether a licensing agreement with a no challenge clause truly excludes third parties depends on whether the underlying patent is perceived to be valid. If the patent is perceived to be valid, third parties are excluded from the market, to the extent that access to the patented technology is essential to market entry. However, what excludes the third parties is not the licensing agreement or the no challenge clause, but the patent—or the perception of the patent—itself. If the patentee does not quantitatively restrict the number of licensees, nothing stops a potential market entrant from reaching a licensing agreement with the patentee and entering the market. If, however, the patent is perceived to be weak, third party entries are restricted to the extent that a third party entrant does not have the economic incentive or the requisite knowledge to challenge the patent. This could be because the litigation costs are prohibitively high in relation to the potential gains from market entry, or the knowledge required to launch a successful challenge can only be gained through commercialization of the technology, which can only take place after a licensing agreement has been reached. Otherwise, third parties would be free to challenge the patent despite the no challenge clause. Meanwhile, a reverse payment agreement will exclude both the potential infringer and third party entrants, especially under the Hatch-Waxman Act.[181]

---

[179] Mungan, *supra* note 171, 30-31.
[180] Miller & Gal, *supra* note 6, at 136, n. 76.
[181] Hovenkamp, Janis & Lemley, *supra* note 170, at 1757.

While it may seem that reverse payment agreements are more anticompetitive because they exclude more rivals, this need not be the case. For example, if the patentee does not produce the final product and only licenses the technology to a licensee on an exclusive basis, there will be only one firm selling products incorporating the patented technology in the market. This arrangement is akin to when a patentee enters into a reverse payment agreement with a potential infringer. In the former case, if the technology grants the patentee monopoly power due to a lack of reasonable substitutes, the monopoly profit will be split through the royalty mechanism. The patentee presumably will only extract part of the monopoly profit through royalty, leaving some monopoly profit to the exclusive licensee. In the latter case, the patentee shares the monopoly profit with the licensee directly through a lump sum payment. In both cases, there is only one producer in the market. Competitive harm is not confined to situations in which the patentee only grants an exclusive license. Even if the patentee grants multiple licenses, she can still maintain its monopoly profit through a variety of licensing practices such as territorial exclusivity, customer exclusivity, or a GE-style price fixing arrangement.[182] Therefore, the number of excluded rivals is a poor proxy for the amount of consumer harm resulting from a patent exploitation practice. What determines whether a particular patent exploitation practice should fall within the ambit of antitrust law should not be the number of excluded rivals, but the amount of possible consumer harm that may result from the practice.

In sum, attempts to distinguish reverse payments and no challenge clauses, and conclude that the latter should be beyond the purview of antitrust law, fail. No challenge clauses can cause consumer harm under certain circumstances. There is no strong justification for excluding no challenge clauses from antitrust scrutiny, as Miller and Gal have argued. One may then wonder why U.S. antitrust law has not addressed no challenge clauses, contrary to the situation in the EU and China. One possible explanation is that the various Courts of Appeals have generally taken a fairly hostile attitude toward no challenge clauses in licensing agreements, notwithstanding the more lenient approach of the Federal Circuit. The Second, Seventh, and Ninth Circuits have by and large held no challenge clauses to be unenforceable as a matter of patent law.[183] Given that the case law suggests that it is usually licensees that challenge the validity of no challenge clauses, it would be more straightforward for the licensee to seek to invalidate the clause under patent law than to attempt to challenge it under antitrust law. This is particularly the case

[182] U.S. v. General Electric Co., 272 U.S. 476 (1926).
[183] See *supra* Section II.A.

given that the Rule of Reason, as opposed to the per se rule, most likely applies.[184] The relative attractiveness of patent law as an avenue for invalidating no challenge clauses probably explains the lack of case law under antitrust law.

B. Probabilistic Patents and Patent Validity

Having established that no challenge clauses should fall within the ambit of antitrust law, it remains to be determined exactly under what circumstances these clauses cause consumer harm A concept highly relevant to the determination of legality of no challenge clauses is the "probabilistic patent." The idea is that unlike real property such as land, where there is much less uncertainty as to the boundary or even the existence of the property right, the validity and scope of a patent are often shrouded in uncertainty.[185] This uncertainty is underscored by statistics that show the failure rate of patentees in defending their patents. Allison and Lemley find that 46% of patents that were litigated to judgment were found to be invalid.[186] A later study found that patentees have their patents invalidated approximately 70% of the time.[187] In the specific context of litigation between generic manufacturers and branded manufacturers, it was found that the patentee loses 48% to 73% of the cases.[188] This is despite the fact that, under the Patent Act, a patentee is entitled to a presumption of validity and a challenger must show by clear and convincing evidence that the patent is invalid.[189] In fact, the success rate is even lower for patent assertion entities, otherwise known as "patent trolls."

---

[184] As the Supreme Court stated in *Topco Associates*, "[i]t is only after considerable experience with certain business relationships that courts classify them as per se violations…" U.S. v. Topco Assoc., 405 U.S. 596, 607-08 (1972). Given that the antitrust courts have had so little experience with no challenge clauses, and that these clauses are not so obviously anticompetitive and lacking in redeeming virtues that they should be summarily condemned, it is highly unlikely that courts would apply the per se rule to no challenge clauses.

[185] The uncertainty of the boundary, or scope, of patent rights is extensively discussed in JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK 46-72 (2008).

[186] John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185, 205 (1998).

[187] Paul M. Janicke & LiLan Ren, *Who Wins Patent Infringement Cases?*, 34 AIPLA Q.J. 1, 20-22 (2006).

[188] SUSAN DESANTI ET AL., FED. TRADE COMM'N, GENERIC DRUG ENTRY PRIOR TO PATENT EXPIRATION: AN FTC STUDY vi (2002); ADAM GREENE & D. DEWEY STEADMAN, PHARMACEUTICALS: ANALYZING LITIGATION SUCCESS RATES 1, RBC CAPITAL MKTS. (2010).

[189] Patent Act, 35 U.S.C. § 282 (2011).

According to one study, patent trolls win only 8% of the cases in which patent validity is litigated to judgment.[190]

A number of commentators have noted the probabilistic nature of patent rights. Pittman notes that "patent validity is an extremely slippery concept. Because the criteria regarding patent validity are so subjective, it is often unclear whether a patent is valid."[191] To underscore the uncertain nature of patent rights, Carl Shapiro famously asserted that a patent does not confer the right to exclude, but only the right to try to exclude.[192] In fact, the U.S. Supreme Court itself expressed skepticism toward the the strength and prevalence of patents.[193]

There are a number of implications from the probabilistic nature of patents. First, a patent is a "bundle of uncertain and imperfect rights,"[194] which are "typically far less valuable than would be idealized 'ironclad' patent rights."[195] This means that patent rights should be calibrated to reflect the probability that a patent will be held valid and infringed, which in turn depends on the scope of the patent.[196] Second, recall that whether no challenge clauses result in consumer harm crucially depends on whether the underlying patent is valid. If the patent is valid, all that a no challenge clause does is to eliminate needless litigation. If, however, the patent is invalid, a no challenge clause may help bolster an invalid patent and preserve the market power and monopoly profit that a patentee does not deserve. Therefore, it would seem that an assessment of the legality of no challenge clauses from an antitrust perspective would require a determination of patent validity. This would introduce a great deal of complexity to antitrust proceedings and would need to be addressed with care. Nevertheless, commentators have discussed the relevance of patent validity to legality under antitrust law at length in the context of reverse payment agreements.[197] Given the apparent similarity between no challenge clauses and reverse payment agreements, this discussion will shed light on how the issue of patent validity should be dealt with in the context of no challenge clauses.

---

[190] John R. Allison et. al., *Patent Quality and Settlement Among Repeat Patent Litigants*, 99 GEO. L.J. 677, 694 (2011).

[191] Pittman, *supra* note 47, at 347-48.

[192] Shapiro, *supra* note 171, at 395.

[193] *See* Dreyfuss & Pope, *supra* note 16, at 972, 980 n. 44 (noting that Supreme Court either reversed or vacated Federal Circuit decisions in favor of patentee between 2006 and 2007).

[194] Shapiro, *supra* note 171, at 395.

[195] *Id.*

[196] Hovenkamp, Janis & Lemley, *supra* note 170, at 1761.

[197] *See, e.g.*, *id.*; Elhauge & Krueger, *supra* note 169; Carrier, *supra* note 170.

C. No Challenge Clauses and Consumer Harm

1. Direct Consumer Harm

No challenge clauses, on their own, do not distort competition or inflict harm on consumers. Only when combined with the right to exclude of a patent and the various competition-distorting licensing practices permissible under patent law do no challenge clauses raise antitrust concerns. As mentioned earlier, if the underlying patent is valid, the patentee is entitled to the supra-competitive prices and the various licensing practices as permitted by patent law.[198] These may result in consumer harm, but this is part of the bargain struck under patent law to sacrifice short-run consumer welfare for long-run dynamic efficiency gains. Antitrust law should accept the implications of the bargain and not intervene. However, if the underlying patent is invalid, then the patentee does not deserve the supra-competitive prices and other profits that may result from the various licensing practices. This is where antitrust law should intervene.

Patent law gives a patentee the right to exclude, or at least the right to try to exclude. This right to exclude, however, need not result in higher prices for consumers if there are reasonable substitutes available in the relevant market. There was a time when the ownership of a patent created a presumption of market power.[199] But that presumption was overturned by the Supreme Court in *Illinois Tool Works, Inc. v. Independent Ink, Inc*. in 2006.[200] This decision stemmed from a realization that where there are reasonable substitutes in the market for the patented product, the patent will not wield market power, and the patentee will not be able to charge supra-competitive prices.[201] And without supra-competitive prices, the no challenge clause will not create consumer harm. Therefore, the patentee possessing market power is a prerequisite for antitrust intervention against no challenge clauses.[202]

The main problem with no challenge clauses is that they prevent licensees from challenging the validity of the patent. The preclusion of licensee challenge

---

[198] HOVENKAMP ET AL., *supra* note 34, at § 1.3.

[199] Int'l Salt Co. v. United States, 332 U. S. 392, 398 (1947).

[200] Ill. Tool Works, Inc. v. Independent Ink, Inc., 547 U.S. 28, 31 (2006) (The question presented to us today is whether the presumption of market power in a patented product should survive as a matter of antitrust law despite its demise in patent law. We conclude that the mere fact that a tying product is patented does not support such a presumption.).

[201] HOVENKAMP ET AL., *supra* note 34, at § 4.2.

[202] Elhauge & Krueger, *supra* note 169, at 323 ("Further, there can be no harm to consumer welfare without market power, which is by definition the power to raise prices above competitive levels.").

would not be of such grave concern if third party challenges were equally probable and likely to succeed. However, there are many reasons to think that licensees are often best positioned to mount a validity challenge, as the Supreme Court noted in *Lear*.[203] Not only do licensees, for myriad reasons, have greater economic incentives to mount a validity challenge, but they also enjoy an advantage in knowledge that increases the likelihood of success of their challenges.[204] First, licensees have more economic incentives than third parties because they are currently paying royalties; which can be avoided if the patent is invalidated. Second, third parties do not have as much incentive to enter the market as licensees because the market is already populated by the existing licensees. The market would be quite competitive by the time they enter, after invalidation of the patent. Third parties also do not have first mover advantage, which gives a competitor cost and marketing advantages over late comers to the market. In a market with homogeneous product and Bertrand competition, even a firm that is equally efficient as existing licensees would not enter the market.[205] Lastly, third party challengers would suffer from the free-rider problem. Each potential third party challenger would want to wait for someone else to shoulder the costs of challenge, because once a patent has been invalidated, it is invalid vis-à-vis all parties.[206] Of course, licensees may also suffer from the same problem. However, licensees have an advantage in overcoming the free-rider problem because they are aware of each other's identity, and can organize more easily to share litigation costs. In contrast, potential third party challengers may not even be aware of each other and may have greater difficulty coordinating.

---

[203] Lear Inc. v. Adkins, 395 U.S. 653, 670 (1969) ("Licensees may often be the only individuals with enough economic incentive to challenge the patentability of an inventor's discovery.").

[204] It should be noted that licensees also suffer from some specific financial disadvantages in mounting validity challenges. Licensees may be subject to the various penalties stipulated in the licensing agreement over validity challenges. And if the agreement provides for automatic termination upon challenge, licensees may be exposed to trebled damages if they continue to produce and the infringement is found to be willful. Given that the licensees have already invested heavily in the production process, it would be very costly for the licensees to cease production. Lastly, licensees have to bear reputation costs, especially as a repeat player in the industry.

[205] Shapiro, *supra* note 171, at 406. This is because the market price post-entry would be driven down to the marginal cost. Given the fixed costs of entry, no entry would occur. Only if the entrant enjoys a significant cost advantage would entry occur. The cost advantage would not only need to enough to allow the entrant to recoup the fixed costs of entry, it would also need to allow the entrant to cover the litigation costs of invalidating the patent.

[206] Blonder-Tongue Labs. v. U. of Ill. Found., 402 U.S. 313 (1971).

Licensees also have an informational advantage over potential third party challengers. They may have gained special knowledge about the patented technology through the license negotiation process,[207] and commercialization of the technology.[208] This is, in no small part, because the licensees will have physical possession of the patented invention, which significantly aids in their understanding of the technology.[209] Licensees will likely also have a good understanding of the prior art based on their experience with the industry in general, and will likely have dealt with similar technology or products in the past.[210] With respect to the specific requirements of patentability, "[a] licensee is likely able to understand, based on its own use, whether the invention falls within the broad scope of patentable subject matter and has a specific and substantial utility. The licensee's use similarly provides a better understanding of whether the patent's written description fully describes the invention and is sufficient to enable one to make and use it without undue experimentation."[211] Therefore, licensees should be better positioned to furnish evidence to challenge patent validity.

If the underlying patent is invalid, no challenge clauses, by precluding licensee challenges, may artificially prolong the exclusion period of a patent, and compel consumers to pay supra-competitive prices for longer than necessary.[212] Therefore, to determine whether a no challenge clause has resulted in consumer harm, one needs to compare the "licensing exclusion period,"[213] that would obtain under the licensing agreement with a no challenge clause, with the "expected exclusion period"[214] that would materialize if the licensing agreement did not contain a no challenge clause.

Under normal circumstances, the licensing exclusion period would be at most the duration of the license, as most no challenge clauses last for the length of

---

[207] Orstavik, *supra* note 7, at 101.

[208] Morris, *supra* note 119, at 220.

[209] Nicholas Roper, *Limiting Unfettered Challenges to Patent Validity, Upholding No-Challenge Clauses in Pre-Litigation Patent Settlements Between Preexisting Parties to a License*, 35 CARDOZO L. REV. 1649, 1675 (2014).

[210] Miller & Gal, *supra* note 6, at 137.

[211] Roper, *supra* note 209, at 1676.

[212] Elhauge & Krueger, *supra* note 169, at 288.

[213] Despite the term, the same concept obviously also applies when the no challenge clause is contained in a settlement agreement. For ease of reference, this Article will use licensing exclusion period to refer to the exclusion period under licensing agreements and settlement agreements. Consent decrees are excluded from this discussion because, as it has been argued earlier, it is inappropriate to apply antitrust scrutiny to an agreement that has been judicially approved.

[214] Elhauge & Krueger, *supra* note 169, at 290.

the license itself. However, it is possible for the no challenge clause to last longer than the length of the license. An example is the licensing agreement in *Bendix Corp v. Balax Inc.*, which prohibited the licensee from ever challenging the validity of the patent, even after the agreement has lapsed.[215] In this case, the maximum licensing exclusion period would be the remainder of the patent term. The situation would be more complicated if the patentee has entered into a range of licensing agreements whose terms vary, or if the patentee, for some reason, has only imposed a no challenge clause on some licenses and not others. In this case, the licensing exclusion period will need to be weighted by the likelihood that a licensee not subject to a no challenge clause will bring a validity challenge. Licensees are not the only source of validity challenges. Unrelated third parties can also launch a validity challenge, which, if successful, will put an end to the patent term and hence the licensing exclusion period. This is likely the major source of uncertainty regarding the licensing exclusion period, as the main reason beyond the parties' control that a license may end prematurely is a third party challenge to the patent. The exclusion period for a settlement agreement would similarly depend on the term of the agreement. If the settlement agreement is meant to remain in force in perpetuity, then the exclusion period would be again the remainder of the patent term plus taking into account third party challenges.

The "expected exclusion period" refers to the exclusion period that would be obtained absent a no challenge clause. Without a no challenge clause, there will be two sources of challenges: the licensees and unrelated third parties. The likelihood that these two groups will bring validity challenges will most likely be different, as explained above. In a world without licenses, the likelihood of a challenge would chiefly depend on the perceived validity of the patent and the resources at the disposal of the potential challengers. Once licensing agreements come into the picture, they affect the economic incentives of parties to bring challenges. When deciding whether to launch a validity challenge, a licensee will compare what she currently earns in the market as opposed to what she would earn in the post-challenge market. One main difference between the two markets is that the licensee would no longer need to pay royalties in the post-challenge market. The state of competition among the licensees may also differ due to current licensing restrictions imposed by the patentee, such as a GE-style price fixing arrangement, output restriction, or territorial exclusivity. Without a valid patent, these restrictions would most likely be illegal and dismantled.[216] A GE-style price fixing arrangement or output restriction would help to maintain supra-competitive prices,

[215] *See* Bendix Corp. v. Balax, Inc., 471 F.2d 149, 153 (7th Cir. 1972).
[216] *See* Hovenkamp, Janis & Lemley, *supra* note 170, at 1746-49.

which benefit the licensees. Territorial exclusivity effectively creates regional monopolies and also benefit the licensees. Without these restrictions, a licensee would have access to the entire market, free from price, output or territorial restrictions imposed by the patentee. However, whether a licensee will stand to gain from such a situation depends on its comparative advantage vis-à-vis other licensees. If a licensee was a more efficient producer of the product due to cost advantages or superior production techniques, she would stand to capture market share from other licensees and would therefore benefit from the dismantling of the license restrictions. But if a licensee was a less efficient producer, she would be better off under the protection of license restrictions, which prevent its competitive disadvantage from being exposed by competitive pressure.

The most obvious difference between the pre-challenge market and the post-challenge market, regardless of the existence of license restrictions, is the entry of third parties. In the pre-challenge market, third parties would be deterred from entering the market to the extent that they are deterred by a perceivably valid patent, or the litigation costs or knowledge requirements of bringing a validity challenge. Once a licensee brings a validity challenge and prevails, the floodgates open for third parties to enter the market. Whether a licensee would achieve a net gain from the removal of the patent depends on its savings from the royalty payment and its competitive advantage vis-à-vis third party entrants. If the licensee was a more efficient producer than the third party entrants, she would worry less about them and would probably achieve a net gain from the removal of the patent. However, if a licensee was a less efficient producer than third party entrants, she would be better off under the existing license restrictions.

The next question is whether third parties would have the same incentives to challenge the patent with and without the no challenge clause. If they do, then the main difference between the licensing exclusion period and the expected exclusion period would be attributed to the licensees. One would think that the third parties' incentive to challenge the patent would be the same with or without the no challenge clause. After all, the no challenge clause does not apply to them, it only affects the licensees. However, it is possible that the no challenge clause will have a signaling effect to potential third party challengers.[217] Such a challenger may think that if all these licensees are willing to accept a no challenge clause, it must

[217] Note, however, that in *Jack Winter, Inc.*, the Court held that the fact that the defendant in that case extended its indemnification program to its customers to cover potential infringement suits from the patentee, even though the defendant had reasons to doubt the validity of the patent, does not render an agreement an unreasonable restraint of trade. Jack Winter, Inc. v. Koratron Co., Inc., 375 F. Supp. 1, 54 (N.D. Cal. 1974).

mean that the licensees are fairly confident that the patent is valid. Otherwise, the licensees would not have agreed to pay royalty and give up their rights to challenge the validity of the patent. This would especially be the case if the licensees are perceived to be firms with intimate knowledge of the technology and would be in the best position to evaluate the validity of the patent. A third party challenger may be deterred from launching a challenge by the fact that a host of knowledgeable firms have willingly accepted a no challenge clause. The third party challenger may be right to put credence in the signaling effect of the licensees' acceptance of the no challenge clause if the licensees have accepted the clause in good faith, after careful examination of the patent. If, however, it turns out that the no challenge clause is the result of a conspiracy between the patentee and the licensees, whereby the licensees would refrain from challenging a highly questionable patent and the patentee will split part of the monopoly profit with the licensees, then the no challenge clause will serve a plainly anticompetitive purpose. Unfortunately, third parties probably cannot distinguish the two situations. Therefore, whether or not the licensees accepted the no challenge clause on good faith, the mere existence of the no challenge clause would somewhat deter a potential third party challenger.

So far, we have only focused on the likelihood of challenges from various sources. A missing piece of the puzzle in determining the licensing exclusion period and the expected exclusion period is the probability that the patent will be upheld when challenged. Formally, the licensing exclusion period should equal the base exclusion period, here the full length of the licensing agreement (and if the no challenge clause prohibits the licensee from ever launching a challenge, it would be the remainder of the patent term),[218] adjusted by the expected invalidity factor, which in turn equals the probability that a third party challenge will be launched times the probability that the challenge will succeed. Let T_{LE} stand for the licensing exclusion period, T_L stand for the duration of the licensing agreement, θ_T stand for the probability of a third party challenge, and θ_{IT} stand for the probability that a third party challenge will succeed. The licensing exclusion period would be represented by:

$$T_{LE} = (1- \theta_T * \theta_{IT}) T_L$$

Likewise, the expected exclusion period needs to take into account the probability that the patent will be held invalid. Formally, the expected exclusion period should equal the duration of the licensing agreement reduced by the expected invalidity factor. This is calculated by multiplying the probability of a third party challenge by the probability that the challenge will succeed, plus the

[218] The same would be true for a settlement agreement without a finite term.

probability of a licensee challenge multiplied by the probability that the challenge will succeed. For reasons discussed previously, licensee challenges may be systematically more likely to succeed than third party challenges,[219] Thus it is important to distinguish them. The base exclusion period in this instance is also the duration of the licensing agreement, and not the full patent term, unless the licensing agreement, or the settlement agreement, lasts for the full term of the patent.[220] Let T_{EE} stand for the expected exclusion period, θ_L stand for the probability of a licensee challenge, and θ_{IL} stand for the probability of success for a licensee challenge. The expected exclusion period is represented by:

$$T_{EE} = [1- (\theta_T * \theta_{IT} + \theta_L * \theta_{IL})] T_L$$

The comparison will be slightly different if the no challenge clause is unenforceable in a particular jurisdiction, as in some of the circuits in the U.S.[221] If the no challenge clause is unenforceable, the licensing exclusion period and the expected exclusion period should in theory be the same, as the licensees are free to challenge patent validity. It would then seem that the no challenge clause inflicts no consumer harm. This would be true if the licensees truly deem themselves not bound by the no challenge clause.[222] It is possible, however, that despite the unenforceability of the no challenge clause, the licensees voluntarily agree not to challenge the validity of the patent, perhaps because the licensees have been offered preferential licensing terms in exchange for a promise not to challenge. In that case, the licensing exclusion period would be the same as if the no challenge clause were binding and enforceable. This would amount to a non-binding agreement by the licensees not to challenge a patent, probably in exchange for some benefit.[223] This possibility has been recognized by commentators. Miller and

---

[219] *See* discussion *supra* Section II.C.1.

[220] This is because the purpose of the exercise is to compare the exclusion period that is affected by the no challenge clause versus the but-for exclusion period. If the full patent term is used as the base exclusion period, then a patentee can easily game the rule by adopting very short licenses, in which case the expected exclusion period will almost always be longer than the licensing exclusion period.

[221] *See, e.g.*, Rates Tech., Inc. v. Speakeasy, Inc., 685 F.3d 163 (2d Cir. 2012); Massillon-Cleveland-Akron Sign Co. v. Golden State Advert. Co., 444 F.2d 425 (9th Cir. 1971); Bendix Corp. v. Balax, Inc., 421 F.2d 809 (7th Cir. 1970).

[222] Even then, there could still be a difference between the licensing exclusion period and the expected exclusion period if third parties are nonetheless deterred by the existence of a no challenge clause from challenging the patent.

[223] While one may argue that a licensee accepting a no challenge clause where such a clause is enforceable is also an agreement by the licensee not to challenge the patent, there is a difference between the two situations. In the former situation, the right to challenge a patent is something that could legally be bargained away. In the latter situation, a bargain over the right to

Gal note that "no-contest clauses may provide a method for parties to cartelize the market based on a patent that was wrongly granted."[224] Hovenkamp, Janis, and Lemley remark that "[t]here is some risk that a patentee may seek to insulate its patent from antitrust challenge by co-opting the most likely challengers with licenses. Where co-option is a problem, the antitrust risks of a settlement are greater than where other potential defendants are likely to challenge the validity of a patent."[225]

Where a no challenge clause is unenforceable, an agreement by the licensees to refrain from challenging the patent would be strongly indicative of a conspiracy between the patentee and the licensees to protect a questionable patent and split the monopoly profit between them, as in the case of illegal reverse payments. One may surmise that the licensees must be generously compensated in order to forego a right that they cannot bargain away under patent law. In addition, one might question why the patentee is willing to offer such generous compensation but for the fact that the patent is of highly questionable validity. The degree of consumer harm inflicted by such a conspiracy, however, would be the same as under an enforceable no challenge clause, because consumer harm is determined by the market power of the patent.[226] Regardless of the enforceability of the no challenge clause, where an agreement not to challenge patent validity exists between the patentee and the licensees, the circumstances under which such a clause would create consumer harm and the degree of consumer harm would be the same. The same analysis applies, but there would be serious reasons to question the validity of the patent.

Based on the foregoing discussion, the following factors should be considered when analyzing whether a no challenge clause results in consumer harm: (1) market power conferred by the patent, (2) the probability of a licensee challenge, which will depend on the licensee's net gain from bringing a challenge, which in turn depends on the market structure, the licensee's comparative advantage, the existence of further licensing restrictions, and third party entrants' comparative advantage, (3) the probability of success for a licensee challenge, (4) the probability of a third party challenge, which may be lowered by the signaling effect of a no challenge clause, and (5) the probability of success for a third party

---

challenge a patent is something that the patent law will not enforce. The licensee will be free to renege on its earlier promise as it sees fit. For the licensee to give up such a right probably requires more generous compensation.

[224] Miller & Gal, *supra* note 6, at 156.

[225] Hovenkamp, Janis & Lemley, *supra* note 170, at 1743.

[226] HOVENKAMP ET AL., *supra* note 34, at § 3.1.

challenge. The probability of a third party challenge and its likelihood of success will be formulated as an affirmative defense in the proposed framework in this Article and will be discussed in Section VI.A. The other three factors will be examined in detail in next Section. If the variables in the two expressions above can be accurately calculated, then a direct comparison between the licensing exclusion period and the expected exclusion period can be made. Otherwise, a qualitative assessment of the various variables will be needed.

2. Impairment of Innovation Incentives

Before moving to a more detailed examination of each of these factors, it is important to discuss some theories of harm that are premised on the impact of no challenge clauses on blocking patents and cumulative innovation. Recall that in the NDRC Guidelines, one of the factors to be considered is whether the underlying patent blocks the implementation of competing patents.[227] Presumably the concern is that if the underlying patent blocks another patent, and the underlying patent is protected by a no challenge clause, exploitation of the blocked patent will be retarded. Obviously if the underlying patent were invalid, then the impediment of the exploitation of the blocked patent would be socially wasteful. However, unless the owner of the blocked patent is itself subject to a no challenge clause, nothing prevents the owner from bringing a validity challenge. If the blocking patent is invalid, the block will be removed. If the underlying patent is found to be valid, then the block stems from the right to exclude of the underlying patent, and not the no challenge clause. The owner of the blocked patent would need to negotiate for a license from the owner of the original patent. If the owner of the blocked patent turns out to be an existing licensee subject to a no challenge clause, one would have expected the licensee to have negotiated for a cross license when entering into the initial licensing agreement. The only scenario in which no challenge clauses would hinder the implementation of a blocked patent is if the licensee subject to a no challenge clause only came up with the technology covered by the blocked patent after entering into the licensing agreement. The no challenge clause would be particularly damaging if the existence of this blocked patent increases the licensee's incentive to challenge the patent, either because the invention process gave him or her new information about the patentability of the blocking patent, or

[227] *Id.* at §34.2 ("Patents may be said to be in a blocking relationship when there is a product or set of products that infringes at least one claim of one party's patent while also infringing at least one claim of another party's patent."); SUZANNE SCOTCHMER, INNOVATION AND INCENTIVES 127-132 (2004) (cumulative innovation refers to innovation that builds on other previous innovation).

the potential to commercialize the blocked patent provides new financial incentives to bring challenges.

The same argument can be made about cumulative innovation. If a cumulative innovation is premised on the underlying patent and cannot be used without a license to the patent, one may be tempted to think that exploitation of the cumulative innovation is retarded when the patent is protected by a no challenge clause. Again, the answer to this argument is that so long as the developer of the cumulative innovation is free to challenge the patent, the no challenge clause should not have any restrictive effect on cumulative innovation. However, that may not always be the case. It is entirely possible, and in fact likely, that the cumulative innovation comes from one of the licensees that developed the improvement during the process of commercialization of the patented technology. However, given that existing licensees already have a license to the patented technology, the innovating licensee should face no obstacles in making use of its improvement (even though the patentee will probably request a license for the improvement), unless the existing license has restricted uses. Therefore, in general, implementation of the cumulative innovation should not be hindered by the existence of a no challenge clause.[228]

D. Factors to Consider When Assessing Consumer Harm

This section examines four factors in determining the likelihood of consumer harm of no challenge clauses. The first factor is what constitutes a no challenge clause for the purpose of antitrust; that is, whether the myriad variations of no challenge clauses should be treated the same in the eyes of antitrust, and whether the kind of agreement which contains the clause alters the analysis. After defining the proper object of analysis, this section moves on to the second factor, market power, which is a prerequisite for consumer harm. A no challenge clause that applies to a patent with no market power will not cause consumer harm. This

[228] Even though no challenge clauses probably will not hinder the deployment of cumulative innovation, there is a scenario in which the emergence of cumulative innovation will increase the licensee's incentive to challenge the original patent. This is if the cumulative innovation is valuable and patentable, but the licensee has serious doubt that the original technology is not patentable. In that case, the licensee can basically take over the exclusivity of the original patentee over the market by patenting its cumulative innovation while invalidating the original patent. The license would thus have significant incentives to challenge, but will be blocked by the no challenge clause. However, consumers are likely to be indifferent between the two scenarios as what will transpire is essentially one patent monopoly replacing another one. If both the original technology and the cumulative innovation are valued by consumers and do not face significant competition in the market, the consumers will have to pay a supra-competitive price either way.

section then examines the factor of patent validity. A no challenge clause will not cause consumer harm if the underlying patent is valid. Lastly, this section analyzes the factor of market structure. Market structure creates different incentives for the licensees to challenge or not to challenge the patent. Licensee incentives matter for two reasons. First, they serve as a proxy for patent validity. A patentee that is unsure about a patent may want to offer the licensees more incentives not to challenge. Second, they tell us how much harm is being done by the no challenge clause; that is,how many potential challenges are being blocked.[229] If no challenge would be forthcoming from the licensees anyway, the no challenge clause would be relatively harmless.

1. Types of Agreements

The first question to consider is whether the analysis should differ based on the type of agreement at issue, be it licensing or settlement, and on the kind of clause at issue, whether it is an outright prohibition, termination-upon-challenge, or other kinds of challenge-penalty clauses.

A number of commentators have correctly observed that there should be no difference between a licensing agreement and a settlement agreement as far as antitrust analysis is concerned.[230] Shapiro observes that "a wide range of commercial arrangements involving intellectual property can be regarded as settlements of intellectual property disputes, either literally or effectively. Virtually every patent license can be viewed as a settlement of a patent dispute: the royalty rate presumably reflects the two parties' strengths and weaknesses in patent litigation in conjunction with the licensee's ability to invent around the patent."[231] While a settlement agreement that is reached after litigation has commenced is clearly consummated in the shadow of ultimate judicial findings on patent validity and infringement, a settlement agreement that is entered into after a dispute has arisen but before litigation has begun likewise falls within the same shadow, albeit a slightly longer one. As Shapiro further notes, "both types of settlements raise the same antitrust issues."[232]

---

[229] If we know the validity of the patent, the amount of potential challenges being blocked would not matter. If we know that the patent is valid, the inquiry would end. However, because we do not know patent validity, every potential challenge is an opportunity to unearth an invalid patent, and by extension consumers' needlessly paying supra-competitive prices. Every potential challenge blocked therefore increases the probabilistic consumer harm of the no challenge clause.

[230] *See* Orstavik, *supra* note 7, at 107; Roper, *supra* note 209, at 1678.

[231] Shapiro, *supra* note 171, at 392.

[232] *Id*.

There is no qualitative difference between a pre-litigation settlement agreement and a licensing agreement, especially one that incorporates a no challenge clause, which suggests that patent validity was within the parties' contemplation and represents an implicit concession of validity on the part of the licensee. The Supreme Court has held in *FTC v. Actavis*, 570 U.S. 756 (2013), that settlement agreements are not immunized from antitrust scrutiny.[233] Recall that the *Rates Technology* court noted that if no challenge clauses in pre-litigation agreements were enforceable, parties could easily circumvent the ban on no challenge clauses in licensing agreements through creative drafting.[234] This is a tacit acknowledgement that the line between pre-litigation settlement agreements and licensing agreements is very thin, if not non-existent.

What about the distinction drawn by some courts concerning the enforceability of no challenge clauses in a settlement agreement that depends on whether discovery on patent merit has taken place? The argument made by those courts is that after discovery on patent merit, the parties have the ability to make a well-informed decision.[235] Presumably, the settling party would not accept a no challenge clause in the settlement agreement if it has grounds to doubt the validity of the patent. If the issue is enforceability of the clause, this argument should carry great weight. However, if the issue is whether the clause is anticompetitive, whether the parties entered into the agreement with full information should not be dispositive. There remains the possibility that the parties have entered into a conspiracy to split the monopoly profit despite both having serious doubts about the validity of the patent, and such an agreement can be anticompetitive even in the absence of a reverse payment.[236] The most that the courts could infer from a post-discovery settlement agreement is that there is a greater probability that the patent is valid.

Whether various kinds of no challenge clauses should be treated differently depends on whether the clause at issue creates a sufficient deterrent to the licensee to mount a challenge. This is because the effectiveness of a no challenge clause is determined by the deterrent effect it creates. Recall that even an outright prohibition in the agreement will only result in damages for breach of contract unless the court enforces it with an injunction.[237] Therefore, most of these clauses

---

[233] *See* Fed. Trade Comm'n v. Actavis, 570 U.S. 756 (2013).

[234] *See* Rates Tech., Inc. v. Speakeasy, Inc., 685 F.3d 163, 171 (2d Cir. 2012).

[235] *See, e.g.*, *Rates Tech.*, 685 F.3d at 171-72; Flex-Foot Inc. v. CRP, Inc., 238 F.3d 1362, 1369-70 (Fed. Cir. 2001).

[236] Elhauge & Krueger, *supra* note 169, at 312-19.

[237] *See* discussion *supra* Section I.B.

ultimately operate on financial incentives, and only differ in degree. That is certainly true of the challenge-penalty clauses. For example, the liquidated damages clause in *Rates Technology* probably provided a powerful deterrent to challenge even though it stopped short of being an outright prohibition.[238] Meanwhile, if the challenge penalty is insubstantial, the deterrent effect will be smaller, and the courts may not want to analyze the clause as an outright prohibition.

The controversy regarding the enforceability of termination-upon-challenge clauses notwithstanding, the practical consequence of termination of a licensing agreement is likely to be coercive enough on a licensee that it functions as an outright prohibition. So long as the licensee has made a substantial investment to commercialize the technology, and has not recouped its investment, the licensee is unlikely to be willing to cease production, which it would be required to do upon termination of the agreement. If the licensee has already recouped its investment and the fixed costs of production are not high, the licensee may be willing to cease production, but probably not for a long period of time. Given that a patent infringement suit can easily last for years, cessation of production is unlikely to be a viable option for most licensees. The alternative would be to keep producing and risk an infringement suit should the patent prove to be valid. If the damages are substantial enough—as they will be if willful infringement is proved—the licensee would only launch a validity challenge if she is highly confident of invalidity.

While it is possible to offer some predictions about the potential coercive effect of some of these challenge-penalty clauses and termination-upon-challenge clauses, in the end, whether a certain clause amounts to an outright no challenge clause will require a case-by-case analysis. This will be the first step in the analysis by a court facing these clauses.

2. Market Power

As explained previously, no challenge clauses will only harm consumers if they allow the owner of an invalid patent to continue to charge supra-competitive prices at the expense of consumers. A patentee will only be able to charge supra-competitive prices if the patent confers market power, which requires there to be few or no reasonable substitutes for the patented product. While determining whether a patent confers market power requires case-by-case analysis, a distinction can be made based on the correspondence between the scope of the patent and product boundary. Patent and antitrust law has long proceeded on the assumption

[238] *See Rates Tech.*, 685 F.3d at 166 (agreement provides for liquidated damages of $12 million).

that one patent results in one product, and hence there is a one-on-one correspondence between patent and product boundary.[239] While that may be true in the pharmaceutical industry, where the final drug product may only incorporate one patent, this is certainly not true in other sectors such as information technology, where products often incorporate hundreds, if not thousands, of patents.[240] The market power analysis will need to be conducted differently in these two scenarios. In the latter case, even if there were no reasonable substitutes for the final product, the market power of that product could not be facilely attributed to a single patent at issue in a case.

In fact, Sidak has gone one step further and argued that no challenge clauses applicable to standard-essential patents ("SEPS") in a patent portfolio are never anticompetitive. This is because the presence of a handful of invalid patents in the portfolio will be inconsequential, and no challenge clauses in such a situation only serve to reduce transaction costs and deter opportunistic behavior by licensees.[241] According to Sidak, the socially optimal number of invalid patents in a portfolio of SEPs is not zero.[242] When the patentee and a licensee negotiate for a license to a portfolio of SEPs, both parties are aware that some of the patents in the portfolio, which may number in the hundreds or the thousands, may be invalid.[243] Parties do not invest the time or the resources to verify the validity of each patent in the portfolio because that would be too costly from a transaction cost perspective.[244] Instead, the parties will assess the value of the portfolio as a whole.[245] The final royalty will reflect the fact that some of the patents may be invalid.[246] Given that the existence of a handful of invalid patents may not make much of a difference to the overall market power of the portfolio,[247] a no challenge clause will not artificially protect the market power of the patentee, and there will be no consumer harm. Meanwhile, allowing licensees to challenge the validity of the patents in the portfolio will give rise to opportunistic behavior:

---

[239] Hovenkamp, Janis & Lemley, *supra* note 170, at 1739.

[240] Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1590-91 (2003).

[241] *See* J. Gregory Sidak, *Evading Portfolio Royalties for Standard-Essential Patents through Validity Challenges*, 39 WORLD COMPETITION (forthcoming 2016), https://www.criterioneconomics.com/docs/evading-portfolio-royalties-for-seps.pdf.

[242] *Id.* at 3.

[243] *Id.* at 10.

[244] *Id.*

[245] *Id.*

[246] *Id.*

[247] *Id.* at 12-13.

> After extensive negotiations, the licensee signs the portfolio license agreement but nonetheless challenges the validity of a few licensed SEPs and refuses to pay the agreed-upon portfolio royalty until the court decides the validity of the challenged SEPs. After the court decides the validity of the few disputed SEPs, the licensee challenges the validity of *another* handful of licensed SEPs and postpones even further its payment of the portfolio royalty. Suppose the licensee repeats this process again and again. That course of action would allow the licensee to postpone its portfolio royalty payments indefinitely and deprive the SEP holder of fair and timely compensation for its innovative contribution.[248]

If this strategy succeeds, the patentee will be denied the royalty it is due. This would impair patentees' innovation incentives in the future, which would be especially damaging for cash-strapped patentees. Therefore, according to Sidak, no challenge clauses in the context of SEP portfolios do not create consumer harm and instead serve the useful purpose of preventing opportunistic behavior.[249] Thus no challenge clauses should be per se legal in the context of SEP portfolios.[250]

There are two key problems in Sidak's arguments. First, in asserting that having a handful of invalid patents is unproblematic, he implicitly assumes that the patents in the portfolio are equally important or valuable, which may not be the case. It is entirely possible for a patent portfolio to be built upon a handful of key patents, and a large number of patents that are either dispensable or can be invented around with relative ease. In such a case, the validity of the key patents would be of central importance to the continual market power of the portfolio. If it turns out that these patents are invalid, the licensees may decide to revoke the license and invent around the remainder of the patents or seek alternatives. Therefore, a categorical statement that having a handful of invalid patents in a portfolio is inconsequential is inaccurate.

Second, Sidak's depiction of the opportunistic behavior by licensees assumes that a licensee can stop paying royalty upon launching a validity challenge. Under existing U.S. case law, it is not at all clear that licensees can stop paying royalties while maintaining the licensing agreement. In *MedImmune Inc. v. Genentech, Inc.*, the Supreme Court explained that in *Lear*, "we rejected the argument that a repudiating licensee must comply with its contract and pay

[248] *Id.* at 14.
[249] *Id.* at 13-15.
[250] *Id.* at 15-16.

royalties until its claim is vindicated in court. 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".[251] In other words, if a licensee repudiates the licensing agreement, she is free to stop paying a royalty, but she will lose the benefit of the agreement if the patent turns out to be valid. If the licensee continues to use the patent during the litigation, the patentee will be able to sue the licensee for infringement and claim damages, including possibly trebled damages. If the patentee chooses not to terminate the agreement upon the cessation of royalty payment, the patentee will be entitled to recover all the royalty accrued during litigation.

This interpretation of *Lear* is echoed by the Federal Circuit in *Cordis Corp. v. Medtronic, Inc.*, in which the court interpreted Lear as saying that while a licensee is free to stop paying royalty during the pendency of a validity challenge, the licensee is not free from facing the consequences of a cessation of payment.[252] If a licensee wishes to continue to invoke the protection of its licensing agreement, she should be required to continue paying royalty to the patentee. If the licensee stops royalty payment, she runs the risk of a breach of contract and liability for damages. Given the importance of these SEP portfolios, it is improbable that the licensee will repudiate the licensing agreement upon challenging the validity of a handful patents in order to save on royalty. The licensee will still need access to the remainder of the patents in the portfolio. Many of the cases in which the courts dealt with the issue of licensor rights and licensee obligations during a validity challenge concerned only a single or a handful of patents.[253] It would be highly unlikely for the courts to hold that a licensee can suspend royalty payment for an entire portfolio of hundreds or thousands of patents simply because she is challenging the validity of a handful of patents. This would be doubly so if the

---

[251] MedImmune, Inc. v. Genetech, Inc., 549 U.S. 118, 124 (2007).

[252] Cordis Corp. v. Medtronic, Inc., 780 F.2d 991, 995-96 (Fed. Cir. 1985). In ruling that a licensee could not avoid facing the consequences of its breach of its license agreement in the course of bringing a patent challenge, the *Cordis* court echoed the view previously expressed in *Morton-Norwich* that "permitting the licensee to unilaterally and offensively ignore his contract obligation to make payments required under the contract, and at the same time denuding the licensor of the remedy of declaring a breach and seeking relief against the licensee as an infringer, . . . does violence to contract principles" in a manner that "might encourage more validity litigation, but at too high a price." Morton Norwich Prods v. Int'l Salt Co., 183 U.S.P.Q. 748, 750 (N.D.N.Y. 1974).

[253] *See, e.g.*, Precision Shooting Equip. Co. v. Allen Archery, Inc. 646 F.2d 313 (7th Cir. 1981); *Cordis Corp.*, 780 F.2d 991; Gen-Probe v. Vysis 359 F.3d 1376 (2004); *MedImmune*, 549 U.S. 118.

courts observe a pattern of repeated challenges over time with the goal of delaying royalty payment.[254]

While a rule of per se legality for no challenge clauses in the context of an SEP portfolio is unwarranted, it remains true that a patent, or a handful of patents, in a portfolio is less likely to wield market power than an entire portfolio of patents. Whether that is indeed the case will need to be determined on a case-by-case basis. Therefore, when facing a no challenge clause applicable to a patent portfolio, the court should first determine whether the portfolio as a whole wields market power. If it does, the court should next examine whether there is a considerable number of patents in the portfolio that are equally important or whether only a handful of patents are important. If the former is the case, then it is unlikely that the no challenge clause will contribute to artificially maintaining the market power of the portfolio, unless the licensee can prove that most of these important patents are of questionable validity. Otherwise, the clause should not be subject to antitrust scrutiny. But if the latter is the case, no challenge clauses could be problematic and antitrust scrutiny should continue.

3. Patent Validity

The legality of no challenge clauses would have been a straightforward question if patents had certain validity and scope. As no challenge clauses would only cause consumers to suffer unnecessary supra-competitive prices if the patent were invalid, and would be perfectly legitimate attempts to eliminate needless litigation if the patent were valid, legality of no challenge clauses would boil down to patent validity. Even though the issue falls under antitrust, the answer must be sought under patent law. Hovenkamp, Janis, and Lemley propound a classification of intellectual property settlement agreements into three categories.[255] The first two categories are relatively straightforward. Unfortunately, no challenge clauses fall within the third category as the competitive harm of these clauses depends on the

[254] *See* Sidak, *supra* note 241, at 15. Sidak's characterization of holders of SEP portfolios as cash-strapped is also unlikely to reflect the reality. Many of the holders of SEP portfolios, at least in the ICT industry where SEPs have the greatest salience, are the likes of Nokia, Ericsson, Motorola, Apple, Samsung, Qualcomm, and Google. These companies are hardly cash-strapped. Moreover, if a company is able to invest in the enormous volume of innovation in order to build an SEP portfolio, it is unlikely to be cash-strapped or otherwise liquidity constrained. It should have no problem borrowing from the capital markets while riding out the opportunistic behavior by licensees.

[255] Hovenkamp, Janis & Lemley, *supra* note 170, at 1720-21. The three categories are (1) agreements that are legal even if the patent is invalid and not infringed, (2) agreements that are illegal even if the patent is valid and infringed, and (3) agreements whose legality depend on the validity of the patent.

validity of the patent. What makes matters more complicated is that the determination of patent validity alone does not answer the question of legality. No challenge clauses are not anticompetitive if the patent is valid, but the opposite is not true. It is not necessarily the case that no challenge clauses are anticompetitive whenever the underlying patent is invalid. Even if the patent is invalid, no challenge clauses would only be anticompetitive if the patent wielded market power, which in turn depends on a host of issues such as the patent-product boundary correspondence and the availability of reasonable substitutes.

This still presents a difficult issue to resolve with respect to patent validity, which is whether it should be assessed from an ex ante perspective at the time the agreement is entered or from an ex post perspective when the no challenge clause is challenged in courts. This is important because if the patentee can demonstrate that the underlying patent is valid, she should be absolved from liability. In the context of reverse payments, most commentators who have expressed a view on this issue have argued that patent validity should be assessed from an ex ante perspective.[256] An ex ante approach makes sense because if patent validity was assessed on an ex post basis, it risks penalizing patentees that in good faith believed its patent was valid and having the courts second-guess the patentee with the benefit of hindsight. This would inflict particular hardship on patentees given the often-unpredictable nature of patent trials.

Having settled on the ex ante perspective, it remains to be decided whether it should be the patentee's subjective perception of patent validity or some objective assessment of what a reasonable patentee would believe in light of the information available to it. The former test is probably easier to administer in that it boils down to simple evidentiary proof of what the patentee herself thought. However, it would be susceptible to abuse. Once patentees know that their contemporaneous statements about patent validity would determine the legality of the no challenge clauses they want to impose, they would inflate their expectations.[257] The latter test will avoid this problem as it does not rely on the party's subjective assessment. It will require an assessment of what a reasonable patentee's belief about validity would be in light of the information at its disposal. For example, if the patent turns

[256] Hovenkamp, Janis, and Lemley assert that reverse payments should be presumptively illegal unless the defendant can show, among other things, that "the ex ante likelihood of prevailing in its infringement lawsuit is significant". *Id.* at 1759. Crane argues the same. Crane, *supra* note 171, at 750. The Department of Justice also shares a similar view. Brief for the United States in Response to the Court's invitation at 10, 22, 28-32, *Ark. Carpenters Health & Wealth Fund*, 604 F.3d 98 (2d Cir. 2010) (Nos. 05-2851-cv(L), 05-2852-cv (CON), 05-2863-cv (CON)), 2009 WL 8385027, at *10, *22, *28-32.

[257] Elhauge & Krueger, *supra* note 169, at 289.

out to be invalid because of anticipation by prior art, which the patentee could not have discovered after reasonable search effort, but otherwise the patent looks valid, it would be reasonable for the patentee to believe ex ante that the patent is valid. This approach would probably entail some inquiry into the objective merits of the patent that Elhauge and Krueger warn against.[258] However, it should avoid a full-fledged trial on the merits of the patent that can be extremely costly and cumbersome.[259]

4. Market Structure at the Licensee Level

The extent of competition among licensees will shed light on the prospect of consumer harm resulting from a no challenge clause. The first issue to confront is whether there is only one exclusive licensee or multiple licensees. This is important for a variety of reasons. First, assuming that the patent provides the patentee with monopoly profit, it would be the easiest for the patentee to share its monopoly profit with one licensee, especially if the patentee does not herself engage in production.[260] The patentee and the licensee will simply negotiate for a split of the monopoly profit by structuring the royalty payment. If, however, there are multiple licensees, and the patentee is unable to restrict competition among them, the monopoly profit can be easily competed away and there would be little to share with the licensees. Therefore, an exclusive license is more conducive to profit sharing between a patentee and a licensee. An exclusive license would be especially suspicious if the no challenge clause is unenforceable under the patent law of the jurisdiction, which means that the agreement is more likely than not a conspiracy between the patentee and the licensee to split profit from a questionable patent.

A related point is whether the patentee also competes in the downstream market; in other words, whether the agreement between the patentee and the licensee is purely vertical or also contains a horizontal element. If the relationship is purely vertical, the licensee may have fewer incentives to challenge the patent,

---

[258] *Id*.

[259] *Id*. at 288.

[260] If the patentee itself also engages in production of the final product, the patentee must be able to coordinate output with the licensee to maintain output at the monopolist level. Otherwise, excess production will pushes prices below the monopolist level. This would not be too difficult to achieve if the patentee produces the key input incorporating the technology itself and the input is used at a fixed ratio to the final product. In that case, the patentee can control the final product output level of the licensee by restricting the supply of the input. If, however, these two fairly restrictive conditions do not hold, the patentee may need to impose some sort of direct restriction on the final output level, which may run afoul of antitrust law.

as she will have less to gain from the patent invalidation. In that case, a no challenge clause may not do much harm. If the licensee and the patentee do compete, the licensee will have greater incentives to challenge the validity of the patent, especially when the licensee does not enjoy a sufficient cost advantage to offset the royalty payment, which means that the licensee's products will always be at a cost disadvantage in the market. Invalidating the patent will thus help to remove this cost disadvantage. In that case, a no challenge clause could preclude a likely challenge. Commentators have gone so far as to argue that it may be a good idea to exempt no challenge clauses in purely vertical relationships.[261] This may be taking the argument too far, as no challenge clauses can produce anticompetitive effects even in purely vertical relationships. Nonetheless, it remains true that the harm of a no challenge clause is likely to be smaller in a purely vertical relationship.

Having multiple licensees means that the post-invalidation market will be more competitive by virtue of the greater number of competitors. It blunts whatever comparative advantage or first-mover advantage that the licensees may enjoy over third party entrants.[262] Licensees will have less to gain in the event of a successful challenge and fewer incentives to mount a validity challenge, resulting in a lower θ_L. Accordingly, the licensing exclusion period is less likely to exceed the expected exclusion period. A no challenge clause in the presence of multiple licensees is hence less likely to result in consumer harm. On a related point, if there are multiple licensees, the patentee may be tempted to forestall competition among them through licensing restrictions such as price, output, or territorial restrictions. These restrictions will help to preserve the monopoly profit and will also increase the loss to a licensee in the event of a successful validity challenge. By bolstering the profit of the licensees, the patentee will discourage a licensee from mounting a validity challenge. This may suggest that the patentee has greater doubts about the validity of the patent and therefore the no challenge clause may be more suspect.

If there are no license restrictions imposed on the licensees and the licensees are highly competitive with each other, then there is less concern about the supra-competitive prices imposed by the patentee. A supra-competitive royalty would only result in higher final product prices for consumers if the market for the licensees was uncompetitive. In a competitive market, the licensees would be forced to absorb the extra cost and would not be able to pass it on to consumers.

---

[261] Orstavik, *supra* note 7, at 108.
[262] Miller & Gal, *supra* note 6, at 149.

Competition among licensees will be particularly keen if the final product is homogeneous and Bertrand competition prevails among them.[263] Consumers will also benefit more from third party entry if the market is characterized by Bertrand competition.[264] In contrast, if the final product is heterogeneous and the market is characterized by Cournot competition, the licensees will be better able to pass the royalty burden on to consumers, and the no challenge clause would be a greater concern.[265]

Moreover, the royalty structure would also have an impact on the extent to which the royalty burden will be passed on to consumers. If the patentee charges a one-time lump-sum payment, the royalty would be a one-time fixed cost that will not be passed on to consumers. It is only when the royalty is charged on a percentage basis of output, revenue, or profit that it constitutes a variable cost that will be possibly passed on to consumers.[266]

Licensee incentives to challenge may be fueled by economic incentives resulting from market structure. As mentioned earlier, licensee incentives to challenge may also be determined by advantages of a technical, cost, or commercial kind of a particular licensee over other licensees and third parties. If a particular licensee has a marked advantage over other licensees, she may benefit more from unfettered competition in a post-invalidation market. Likewise, if a licensee has an advantage over third parties, she may be less deterred by the

---

[263] Daniel F. Spulber, *Bertrand Competition When Rivals' Costs Are Unknown*, 43 J. INDUS. ECON. 1 (1995).

[264] *See* Shapiro, *supra* note 171, at 401. Bertrand competition refers to an oligopolistic market in which firms produce homogenous product and compete on price. Competition will eventually drive the price to the level of marginal cost and firms earn no supra-competitive profit. DAVID BESANKO & RONALD R. BRAEUTIGAM, MICROECONOMICS 533-34 (4th ed. 2010).

[265] Cournot competition refers to an oligopolistic market in which firms produce differentiated products and compete on output level. Prices will exceed marginal cost and firms will exhibit some market power. BESANKO & BRAEUTIGAM, *supra* note 264, at 535-40.

[266] A final remark is in order. The foregoing discussion may seem to exhibit a degree of inconsistency in that in some instances, a reduced incentive to challenge as a result in market conditions points to legality, while in some other instances it suggests illegality. The reason this is the case is because one needs to look deeper to see what are the reasons for the reduced incentive. If the reduced incentive is due to sharing or bolstering of monopoly profit by the patentee, for example, by way of favorable royalty or license restrictions, it suggests that the patentee has serious doubts about the validity of the patent and needs to induce licensees to accede to a no challenge clause with increased profit. In that case, reduced incentive points to illegality. If, however, the reduced incentive is due to competitive pressure in the market, for example, because of the existence of multiple licensees or a homogeneous product in the market, then reduced incentive is a positive indication that the no challenge clause does not foreclose likely challenges. In that case, reduced incentive to challenge suggests legality.

prospect of post-invalidation third party challenges from mounting a validity challenge. In both circumstances, the licensee will have strong incentives to challenge, and a no challenge clause will be harmful by blocking likely challenges. A third party may also enjoy an advantage in technical know-how, cost, or commercial attributes. To the extent that a third party enjoys such an advantage, it will be more likely to challenge the patent, in which case the no challenge clause will not inflict much harm by precluding licensee challenges.

IV
JUSTIFICATIONS FOR NO CHALLENGE CLAUSES

A host of arguments, some based on innovation incentives, and some on transaction costs, have been offered to justify the upholding of no challenge clauses across the board. This section examines these arguments and concludes that none of them justifies a blanket approval of no challenge clause, regardless of potential harm to consumers.

A. Dynamic Efficiency Justifications

1. Prohibiting No Challenge Clauses Would Cause Patentees Not to License the Patent

Some argue that if no challenge clauses were prohibited, patentees would be discouraged from engaging in licensing at all, which would have adverse welfare consequences.[267] It is widely agreed that licensing can be welfare enhancing.[268] A patentee may choose to license its technology to a third party producer because that producer has lower costs of production, a better distribution network, or an otherwise superior ability to commercialize the product.[269] If the patentee is forced to forego this option, she may do one of the following three things. First, she may choose to commercialize the product on her own, even though she may have to do so at higher costs. Second, she may have to choose an inferior downstream producer which may for one reason or another be less likely to challenge the patent. And the comparative advantage of licensees and their incentives to challenge are often correlated.[270] Lastly, if the technology is difficult to reverse

[267] *See* Miller & Gal, *supra* note 6, at 148, Taylor, *supra* note 22, at 233.

[268] Morris, *supra* note 119, at 229.

[269] *See id.*

[270] Miller & Gal, *supra* note 6, at 148 (The patentee may choose to license the patent to a less efficient licensee "that does not have the resources, stamina, or knowledge to challenge the patent. In particular, P may choose a firm that lacks a significant comparative advantage over other potential producers, since it is often the comparative advantage that propels a licensee to challenge the patent and capitalize on first-mover status.").

engineer, she may choose to rely on trade secret protection instead. Society would be worse off in this case because without the disclosure mandated by the patent system, it may never benefit from the knowledge following from the invention.[271]

In a regime where no challenge clauses are prohibited if the patent is of questionable validity ex ante, prior to making the decision to license its invention, the patentee will perform the following analysis. She will examine its patent closely and decide how strong it is. If she believes that the patent is strong, she first would have much less to worry about from validity challenges. In any case, the legal framework proposed in the previous section would allow this patentee to impose a no challenge clause, so long as the belief in patent validity is reasonable. The patentee can safely proceed to license its patent. If she believes that the patent is weak, she will have to decide whether she wants to take a chance, especially when she knows that licensees will learn more about the technology from the commercialization process and be in a better position to challenge. Patentees are uniquely placed to evaluate their own patents as they may have access to unique information pertaining to patent validity.[272] They should be able to make this choice in a very well informed manner. If we allow the owner of a weak patent to impose a no challenge clause, the patentee will be able to enjoy the double benefit of profiting from a patent of questionable validity and doing so in the most cost-effective manner by deploying the most efficient downstream producers. There are good reasons to question whether owners of such a patent should be afforded such an advantage.

A related concern is whether it is fair to allow licensees to use information provided by the patentee or otherwise gleaned from the commercialization process to launch a validity challenge. A possible objection is that it is unfair to allow the licensees to use information provided by the patentee against the patentee. There are two reasons that this should not be a serious concern. First, the extent to which this is a concern is inversely related to the strength of a patent. The owner of a strong patent is unlikely to be too worried about such a challenge. Second, the patentee has control over what information to disclose, so she could presumably choose to hide incriminating information from the licensees. If it turns out that it is impossible to impart sufficient technical knowledge to the licensees to commercialize the technology without also revealing incriminating information, it would suggest that the flaw in the patent is quite fundamental. One wonders

---

[271] *Id.*

[272] Dreyfuss & Pope, *supra* note 16, at 991. For example, only the patentee is likely to know the true date of invention and the identity of the inventor.

whether the law should provide assistance to facilitate the continual validity of such a patent.

2. Prohibiting No Challenge Clauses Would Allow Patentees to Protect Themselves from Erroneous Invalidation by the Courts

Due to the probabilistic nature of patents, patents may be erroneously struck down by the courts and therefore patentees are entitled to use no challenge clauses to guard against that risk. Some say that it is in the "interest" of patentees to shield their patents from challenges.[273] Others assert that "[a]n increased rate of challenges to patent applications might create a super-optimal number of false negatives … That, in turn, could lead to sub-optimal investment in innovation."[274] Of course it is in the interests of patentees to protect their patent rights, just like it is in the interests of competitors to fix prices. However, the question is whether this is an interest worthy of protection under antitrust law. These arguments betray a fundamental misunderstanding of the probabilistic nature of patent rights. These rights are uncertain because of the various requirements of patentability, which do not lend themselves to application with mathematical precision, and because of the need to delineate the boundary of a patent by interpreting its claims, which has an element of subjectivity like every other interpretation exercise. Inherent in this uncertainty are two elements: (1) the same fact may be subject to different, but equally reasonable and valid, interpretations that may produce different results, (2) and the conclusion may simply be wrong. For example, reasonable people may differ on whether a particular invention is novel enough to warrant patent protection, which falls within the first kind of uncertainty. Meanwhile, whether the invention was in public use more than a year before the date of the application falls within the second kind of uncertainty, for which there is usually a definitive answer. In the ideal world, we would like to eliminate the second source of uncertainty while preserving the first, for the first source of uncertainty is inherent in the nature of patents. However, no one has devised a mechanism that will allow us to do that. No challenge clauses remove the second source of uncertainty, but unfortunately also eliminate the first.

The elimination of the first source of uncertainty results in overcompensation for the patentees. Patentees are entitled to rewards in the form of a royalty or supra-competitive prices. However, this reward should be adjusted by the probability that the patent will be held invalid. Patentees were never meant to

[273] Orstavik, *supra* note 7, at 101.
[274] Crane, *supra* note 171, at 756.

be entitled to receive a risk-free reward for their inventions.[275] As Hovenkamp, Janis, and Lemley argue, "[a]ssertions that patentees are entitled to treat their patents as free from uncertainty, or that they will not receive the proper incentives unless allowed to exclude competitors on the basis of dubious patents, simply misunderstand the structure of the patent system."[276] The corollary of the assertion that patentee reward is by nature risk-adjusted, is that patentees are not entitled to expunge the risk in its reward by imposing no challenge clauses.

Moreover, susceptibility to the second source of uncertainty is not unique to patents. Every property right or other kinds of economic entitlement that require court adjudication are susceptible to false negatives. No one has argued that owners of these economic entitlements, such as contractual rights, should be allowed to shield themselves from erroneous adjudication by barring legal challenges. There are no obvious reasons that patentees should be given special treatment.

3. Prohibiting No Challenge Clauses Would Reduce Incentives to Innovate

Another argument related to the one described immediately above is that prohibiting no challenge clauses and allowing licensees to mount indiscriminate challenges would bring uncertainty to patentee reward and undermine innovation incentives.[277] For the quid pro quo underlying the patent system to function properly, "inventors need to be confident that their rights will be secure. Such confidence is fundamental to providing this incentive to innovate."[278] The same argument used to dismiss the concern about false negatives in patent validity applies with equal force here. Patentees should only be entitled to risk-adjusted reward for their invention.[279] So long as they receive such a reward, innovation incentives will be properly maintained. In fact, if all patentees, regardless of the strength of their patents, receive a risk-free reward, there will be over-compensation for weak patents.[280] Weaker patents may be less novel and have less technical merit that deserves less compensation from society. Since weak patents are presumably easier and less costly to come up with (perhaps because they are less novel or more obvious), potential inventors will rationally gravitate toward investing in inventions of more questionable merit. Because no challenge clauses

[275] Shapiro, *supra* note 171, at 435.

[276] Hovenkamp, Janis & Lemley, *supra* note 170, at 1761.

[277] Crane similarly argued that prohibiting reverse payments would undermine innovation incentives. *See* Crane, *supra* note 171, at 751.

[278] Roper, *supra* note 209, at 1662.

[279] Shapiro, *supra* note 171, at 395.

[280] Elhauge & Krueger, *supra* note 169, at 294.

produce greater consumer harm when the patent is weak, the concern that the prohibition of no challenge clauses will hamper innovation incentives is unfounded.

Other scholars raise a slightly different argument. No challenge clauses will not result in invalidation of questionable patents but will simply result in less return to patentees.[281] They describe the chain of events as follows. "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."[282] At first glance, this argument appears to undermine the premise that at least under some circumstances, prohibiting no challenge clauses will result in invalidation of questionable patents. Upon closer analysis, however, these arguments are unsupported. This argument poses two issues: the wealth transfer from patentees to licensees and the lack of invalidation of questionable patents. In regards to the first issue, wealth transfers between patentees and licensees are of no concern to antitrust law as long as they do not inflict harm on consumers. Wealth transfer from the patentee presumably will reduce innovation incentives, and the decrease in innovation incentives argument has been addressed above. In regards to the second issue, the lack of invalidation of questionable patents is of course a serious concern, but the problem is overstated. First, if the patent is so weak, it is unlikely that the patentee can recruit the licensee to settle the invalidity suit without offering some substantial financial incentives,[283] such as reverse payments. So long as reverse payments are carefully scrutinized by the Agencies and the courts—as they are—the patentee will have limited ability to settle these suits. Second, one would expect that if patentees are constantly forced to share a substantial portion of their surplus to get a licensee to agree to settle the invalidity suit, the return for investing in such weak patents will decrease over time and fewer and fewer patentees will pursue these patents.[284] This is likely to be beneficial to society in the long run.

---

[281] Dreyfuss & Pope, *supra* note 16, at 974.

[282] *Id.*

[283] How substantial the financial incentives will be will depend on how lucrative the market is and how much market power the patent wields.

[284] Elhauge & Krueger, *supra* note 169, at 294.

B. Transaction Costs-Based Justifications

1. No Challenge Clauses Protect Patentees from Wasteful and Vexatious Litigation

One commonly invoked justification for no challenge clauses is that they protect patentees from wasteful and vexatious lawsuits from licensees.[285] Litigation costs are transaction costs in the patent system that can be avoided by no challenge clauses. There are two layers in this argument. First, the litigation costs incurred in these invalidity suits are wasteful and to be avoided if possible. Second, patentees may sometimes need no challenge clauses to protect themselves from aggressive or bad faith licensees. It is a common misconception that patentees must hold greater bargaining power than the licensees, and patentees coerce the licensees to accept oppressive contract terms. For example, in *MedImmune*, the Supreme Court "mistakenly characterized a licensing situation as inherently 'coercive' and akin to government regulatory action." [286] On the contrary, patentees that are thinly capitalized or in emerging sectors such as biotechnology may suffer from power asymmetry problems and have poor bargaining power.[287] Occasionally, there are bad faith licensees who enter into a licensing agreement simply to challenge patent validity.[288] According to Alfaro, "licensees will [after MedImmune] seek to enter into license agreements in an attempt to cap their exposure to infringement liability and then seek a declaratory judgment on the validity of the patent in an attempt to avoid that exposure altogether."[289]

Although an invalidity suit brings about litigations costs, these costs are not necessarily wasteful. As has been acknowledged in this article, if it was known a priori that the patent is valid, then invalidity suits would indeed be needless and wasteful, and no challenge clauses would serve a useful function. However, there is no way to know the validity of a patent until it has been adjudicated in court. Therefore, such litigation expenses are the necessary consequence of probabilistic patents and the imperfections of the patent examination system administered by the Patent and Trademark Office.[290] Moreover, not every patent will be examined judicially. Only patents that are economically valuable will be so examined. And if a patent is valuable—perhaps because it commands market power—then it may not

[285] Taylor, *supra* note 22, at 216, 227.
[286] Ritchie, *supra* note 14, at 142.
[287] Dreyfuss & Pope, *supra* note 16, at 988.
[288] Taylor, *supra* note 22, at 216.
[289] Alfaro, *supra* note 32, at 1304.
[290] Allison & Lemley, *supra* note 186, at 205.

be a bad idea for society to expend some resources to ensure that the supra-competitive prices are not borne in vain.

As for protection from bad faith or aggressive licensees, one cannot argue against protection from vexatious lawsuits, especially for patentees that have poor bargaining power or are otherwise unable to protect themselves. Such lawsuits serve no useful purpose in society and merely incur needless costs. Patentees, weak or strong, deserve protection from them. However, an invalidity lawsuit is only groundless if the validity of the patent is strong. For patents of questionable validity, such suits can no longer be called vexatious. They do in fact serve a useful social purpose. The framework proposed in this article concedes that no challenge clauses would be permissible if a reasonable patentee, ex ante, would believe that the patent is valid in light of the information available to her. Therefore, no challenge clauses would only be subject to antitrust scrutiny if the patent was of questionable validity ex ante. For such patents, invalidity suits would no longer be vexatious, and patentees should then not be shielded from them.

2. Prohibiting No Challenge Clauses Will Reduce Licensees' Incentive to Scrutinize Patent Validity before Entering into a Licensing Agreement

It has also been argued that if no challenge clauses are allowed and enforced, they will merely force licensees to undertake careful scrutiny of the patents in advance and bring up any dispute before they enter into a licensing agreement.[291] Such validity disputes will be more easily settled when the licensees have not invested in commercializing the technology, and the patentee and the licensee are not entangled in a licensing relationship. Licensees also avoid paying unnecessary royalties. Society will be better off if invalid patents are weeded out early.

This argument would be valid if licensees had the same knowledge and economic incentives to challenge the patent before and after entering into the licensing agreement, but they do not. Pre-licensing evaluations of patent validity are unlikely to be perfect due to licensee inability to acquire information uniquely in the hands of the patentee and perhaps also third parties.[292] Licensees may acquire important information about the technology in the process of commercialization. The informational advantages of licensees over third parties discussed previously would put licensees in a much better position to challenge the patent after, as opposed to before the agreement is reached.[293] Licensees may also want to hold off mounting a challenge until they have acquired a first mover advantage in the

[291] Dreyfuss & Pope, *supra* note 16, at 988.
[292] *See supra* note 269.
[293] *See* discussion *supra* Section I.A.

market.[294] After the patent is invalidated, the market will likely become competitive, and a licensee with a first mover advantage will have a significant edge over potential new entrants. Recall that in a market with homogeneous product and Bertrand competition, even a firm that is equally efficient as existing licensees would not enter the market. Therefore, pre-licensing challenges are no substitute for post-licensing challenges.

C. Other Justifications

1. Prohibiting No Challenge Clauses Will Result in Higher Royalty

A number of commentators have argued that, contrary to intuitive belief, prohibiting no challenge clauses would not only fail to bring down prices for consumers, but could result in higher royalties, which could result in higher prices for consumers.[295] Their argument is that when a patentee can impose a no challenge clause, she is more assured that the patent will not be challenged, and its royalty income stream will continue.[296] However, when the no challenge clause is deemed unlawful, the patentee would charge a higher royalty for the higher risk she now assumes. This would be a very damaging charge, as it essentially predicts that antitrust regulation of no challenge clauses would be counterproductive.

This argument should not shield no challenge clauses from antitrust scrutiny for two reasons. First, not every patentee can pass on the extra risks she assumes in the form of higher royalty. This depends on the bargaining power of the patentee, which in turn depends on its market power.[297] Not every patentee wields market power ,so not every patentee will be able to pass on the extra risks to the licensees. Second, and more importantly, the effect of antitrust scrutiny should be evaluated from an aggregate perspective. In some instances, consumers may have to pay higher royalty indirectly through higher product prices when certain patentees cannot employ no challenge clauses. While in other instances, they would benefit when the absence of no challenge clauses allows licensees to challenge invalid patents and royalties are removed for technologies covered by these patents. Prices for products incorporating these technologies would presumably drop. It is difficult to conclude a priori that consumers are necessarily worse off in this new state of affairs. It is entirely possible that consumers may benefit more from the elimination of royalties in some products while suffering from slightly higher

---

[294] Miller & Gal, *supra* note 6, at 153.

[295] *Id.* at 150; Taylor, *supra* note 22, at 233; Brenner, *supra* note 47, at 58, 66-68.

[296] Miller & Gal, *supra* note 6, at 150; Taylor, *supra* note 22, at 233; Brenner, *supra* note 47, at 58, 66-68.

[297] Dreyfuss & Pope, *supra* note 16, at 987.

royalties in other products, especially when the patentee is unable to pass on the full costs of the extra risks to the licensees. Therefore, it is inaccurate to say that antitrust regulation of no challenge clauses would be counterproductive.

V
PROPOSED ANALYTICAL FRAMEWORK

A. The Proposed Framework

In light of the foregoing discussion, this article proposes a new analytical framework for determining the legality of no challenge clauses. However, first it is important to clarify the implications of the algebraic expressions for the licensing exclusion period and the expected exclusion period. As explained previously, whether a no challenge clause inflicts consumer harm entails a comparison of two periods. If the licensing exclusion period exceeds the expected exclusion period, the no challenge clause should be deemed illegal. By simply looking at the two expressions, it would be obvious that under almost all circumstances, the licensing exclusion period would be longer than the expected exclusion period due to the presence of the term ($\theta_L * \theta_{IL}$) in the expression for the expected exclusion period. The two periods would only be the same if this term equals zero, which is highly unlikely if there is a more than negligible chance that licensees would bring successful challenges. The implication would be that no challenge clauses should be illegal across the board.

Two clarifications are in order. First, no challenge clauses would only be an antitrust concern if the patent wields market power. Therefore, the comparison between these two expressions would not be undertaken at all for many no challenge clauses. Second, it is highly unlikely that an actual calculation of these two periods will be attempted in judicial proceedings. Many of the variables in these two expressions are difficult to ascertain and any offer of estimation results for these variables is likely to be challenged by the opposing party. In practice, judicial proceedings will likely boil down to a qualitative evaluation of the various indicators of the probability and likelihood of success of challenges and hence consumer harm. No challenge clauses are unlikely to be condemned unless they substantially increase the licensing exclusion period above the expected exclusion period.

Four factors have been enumerated for consideration in determining the likelihood of consumer harm: types of agreement, market power, patent validity, and market structure at the licensee level. Two of them can be dealt with here. For the purpose of the legality of no challenge clauses, no distinction should be drawn between licensing agreements and settlement agreements. Meanwhile, consent

decrees should be beyond the reach of antitrust law. Whether the myriad variations of no challenge clauses should be treated the same depends on the deterrent effect on challenges wrought by the clause at issue. This would entail a case-by-case analysis. The foregoing discussion should make it clear that, unsurprisingly, no challenge clauses should be subject to the Rule of Reason, and not a per se treatment.[298] The extent of consumer harm of these clauses is so circumstance-specific that an individual examination is necessary. One cannot categorically say in advance that they are always legal or illegal. And as no challenge clauses will not create consumer harm absent market power, the first part of the analytical framework would be a market power screen.[299] If the plaintiff fails to show that the patent at issue wields market power, the case should be dismissed right away.[300]

Another feature to be considered early in the analysis is whether the no challenge clause is enforceable in the jurisdiction at issue, to the extent that it can be ascertained. This is relevant because if no challenge clauses are unenforceable, but licensees still voluntarily agree to abide by them, it is valid to question why the licensee would do so. However, it may not be possible to tell whether a licensee refrains from challenging because of a conspiracy between itself and the patentee or because the licensee genuinely believes that the patent is valid. It then becomes important to examine whether the licensee has been offered unusually generous licensing terms or other financial incentives. One obvious reason that the patentee will offer such generous incentives is because she wants to protect a weak patent.

---

[298] HOVENKAMP, *supra* note 34, at § 5.6.

[299] Recall that under the NDRC draft guidelines, no challenge clauses are treated as both a potential restrictive agreement and an abuse of dominance. Transposed to the U.S., it means they would fall under both Section 1 and Section 2 of the Sherman Act. Given the necessary existence of an agreement, no challenge clauses would definitely fall under Section 1. No challenge clauses could probably qualify as a monopolization offense under Section 2 provided monopoly power. However, given the substantially higher market power threshold under Section 2, in reality, it is likely that most plaintiffs would invoke Section 1 rather than Section 2, just as in the case of tying and exclusive dealing (although in the case of the latter, Section 3 of the Clayton Act is also invoked).

[300] The application of this market power screen would be more complicated if the case concerns a portfolio patents, and as is often the case, a portfolio of SEPs. If the case concerns an SEP portfolio, on one level it is simpler because market power is obvious. On another level, it is more complicated because no challenge clauses will only have an anticompetitive potential with respect to a patent portfolio if it can be shown that there are a small number of highly important patents in the portfolio while the rest are subsidiary. Otherwise, if the patents in the portfolio are of roughly equal importance, Sidak was right that the invalidation of a handful of patents would not make much of a difference. Therefore, if the case concerns a patent portfolio, it is incumbent on the plaintiff to show that the invalidation of a small number of patents will have a decisive impact on the amount of market power wielded by the portfolio.

Therefore, the unenforceability of the clause together with unusually generous financial incentives could be viewed as a strong indication of likely consumer harm. Of course, if the agreement is accompanied by generous reverse payments by the patentee to the licensee, it should attract antitrust scrutiny, as it would under *Actavis*.[301] However, even in the absence of a reverse payment, the licensee could still be offered an advantageous royalty rate, which may indicate an attempt by the patentee to recruit the licensee into a conspiracy.

Next, the defendant should be given the opportunity to offer two affirmative defenses. The first is based on the notion that no challenge clauses will not inflict harm on consumers if the patent turns out to be valid because the clauses only serve to avoid needless litigation expenses. Therefore, if the defendant can prove that a reasonable patentee with the information accessible to the defendant ex ante would believe that the patent is valid, the case should be dismissed. The second is premised on the idea that a no challenge clause would not help to defend an invalid patent if in spite of it, a third party with a similar level of knowledge as the licensees was likely to launch a validity challenge. This defense is necessary because the remainder of the analysis focuses on the licensee challenges. This is the only place in the analytical framework where third parties are considered. This defense has two components: the probability of a third party challenge and the probability of success of a third party challenge. Recall from previous discussion that two of the factors that need to be considered in analyzing the extent of consumer harm of no challenge clauses are the probability of a third party challenge, and the probability of success for a third party challenge.[302] These factors are the focus of this second affirmative defense.

Determining the probability of a third party challenge necessitates a consideration of the signaling effect of no challenge clauses in licensing agreements; that is, the extent to which third parties are deterred from mounting validity challenges due to the existence of no challenge clauses. Another relevant factor is whether a third party enjoys technical, cost, or commercial advantages over the licensees and other third parties that it would be likely to bring a challenge. The proof of a similar level of knowledge is important because licensees usually enjoy informational advantages over third parties in mounting validity challenges. The level of knowledge serves as a proxy for the probability of success of a third party challenge. The probability of third party challenge and the probability of success required can be adjusted depending on the magnitude of probable consumer harm. If the patent wields a substantial degree of market power

[301] Fed. Trade Comm'n v. Actavis, 570 U.S. 756 (2013).

[302] There was also market power, which has already been dealt with.

or perhaps even is monopolistic, the courts should only let a defendant off if a third party challenge is highly probable and likely to succeed. If the magnitude of probable consumer harm is lower, showing a lower degree of likelihood and probability of success may suffice. This defense admittedly may be difficult to establish given the knowledge requirement. However, this requirement is necessary to ensure that the challenge that may materialize would be effective.

If the plaintiff manages to prove market power and the two affirmative defenses are unavailable, then the analysis moves to a holistic assessment of the relationship between the patentee and the licensees and the market structure at the licensee level. These are relevant for determining what kind of economic incentives are present, which determines the probability of a licensee challenge.[303] As discussed earlier, economic incentives are relevant because they tell us both how hard the patentee is trying to entice the licensees not to challenge and how much harm is being done by the no challenge clause as indicated by how likely challenge is being blocked. If the assessment shows that the market structure is such that licensees have substantial incentives to challenge, then the no challenge clause will be blocking a probable challenge. Likewise, if there are indications that the patentee is offering licensees substantial financial incentives, then there are grounds to question the validity of the patent, which means that the no challenge clauses do artificially extend the period during which consumers are saddled with supra-competitive prices. In both cases, the no challenge clause should be struck down. One final aspect of the market structure to consider is whether competition is keen at the licensee level. If it is so, then the licensees would have limited ability to pass on the royalty payment onto the final consumers, in which case consumers suffer little harm even if the patent turns out to be invalid. Lastly, one also needs to consider whether a licensee enjoys technical, cost or commercial advantages over other licensees such that she has strong incentives to bring a challenge. If such a licensee exists, a no challenge clause will be more damaging.

Finally, to account for the competitive harm of no challenge clauses in hindering the exploitation of a blocked patent, the plaintiff should be required to show that the blocked patent is commercially valuable and was only created after the second inventor had entered into a licensing agreement with a no challenge

[303] (2) is not examined in this analytical framework because whatever the probability of success for licensee challenges is, it represents the best chance we have to invalidate the potentially invalid patent. It is important to consider the probability of third party success because we need to ensure that third party challenges can serve as a substitute for licensee challenges. But given the licensee challenges are the best chance we have, the analysis will simply have to take it as a given depending the factual circumstances of each case.

clause with the initial patentee. The proof of commercial value is important because blocking a patent of little commercial value will not result in significant harm to consumers.

B. An Evaluation of the Approaches Taken by the Three Jurisdictions

It should be obvious that none of the three jurisdictions surveyed currently take a suitable approach to no challenge clauses. To the extent the lack of case law on these clauses in the U.S. is a reflection of per se legality under antitrust law, the current standard is clearly too lenient. It has been amply illustrated in this article that no challenge clauses can cause consumer harm under certain, albeit somewhat narrow, circumstances. Therefore, a per se legality approach is unwarranted.

Meanwhile, the EU approaches (to the extent that the Commission's approach under the TTBER differs from the court's approach) in some places are fairly consistent with the analysis in this article. Arguably the part of the EU jurisprudence on no challenge clauses most consistent with the reasoning in this article is *Bayer v. Süllhöfer*, in which the ECJ declared that there is no reason to treat licensing and settlement agreements differently and a determination of the legality of these clauses entails an examination of the legal and economic contexts.[304] This is consistent with the general direction of this article's framework that there needs to be a Rule of Reason-type analysis of the economic incentives created by the market structure. The European Commission's approach is stricter than warranted by the economics of no challenge clauses. There are no good reasons to treat no challenge clauses as an excluded restriction, at least not without a proof of market power. The Commission's continual permissive attitude toward no challenge clauses in settlement agreements is unjustified, but its increasingly cautious attitude is a step in the right direction. Meanwhile, by saying that the focus is on whether the loss of profit would serve as a sufficient deterrent to licensee challenges, the Commission was spot on in its analysis of termination-upon-challenge clauses.

To the extent that the reference to "eliminate or restrict competition" in Article 10 of the SAIC Regulation entails a detailed analysis of competitive effects and market conditions, the SAIC's approach would be consistent with the analysis in this article.[305] The NDRC has manifested two approaches to no challenge clauses, one in its draft guidelines and the other in the *Qualcomm* decision. The enumeration of a number of factors to be considered in Article 2(1)(3) of the draft

[304] Case 65/86, Bayer AG v. Süllhöfer, 1988 E.C.R. 5249.

[305] Provisions on Prohibiting the Abuse of Intellectual Property Rights to Exclude and Restrain Competition, *supra* note 53, at art. 10.

guidelines is a step in the right direction, especially in comparison to the practically per se approach taken in the *Qualcomm* case. However, the factors listed are not exactly the relevant factors and fail to focus on the economic incentives of the licensees to challenge, which may depend on the market structure on the licensee level among other factors. The NDRC approach taken in the *Qualcomm* case is clearly problematic in light of the foregoing analysis. The NDRC's focus on the right of the licensees to challenge is misguided. And the complete lack of attention to the fact that the no challenge clauses in that case were applied to an SEP portfolio may have overstated the practical impact of these clauses. It is entirely possible that the invalidity of a handful of patents in Qualcomm's portfolio would make no difference to its market power. In that case, the claim should have been dismissed.

CONCLUSION

This article examines a patent licensing practice that has long escaped the attention of U.S. antitrust law. It analyzes the circumstances under which no challenge clauses can cause consumer harm and rejects the approach prevailing in Europe and in U.S. patent law that no challenge clauses are harmful because they frustrate the public policy of the removal of invalid patents. If that were the main policy consideration behind the analysis of no challenge clauses, it would result in a per se rule against them, which would be unduly harsh. Instead, it proposes an analytical framework that would allow the courts to approach these clauses in a systematic manner, giving due regard to the fact that many no challenge clauses do not pose consumer harm. The framework provides for a number of devices and defenses to help screen out cases in which consumer harm is unlikely, so that courts only need to get to the complex Rule of Reason analysis when it is truly necessary and likely to be fruitful. This article further explains how different elements in the market structure should illuminate the analysis and makes the important observation that market structure is relevant because it affects the economic incentives of licensees to challenge the patent. Importantly, this article rejects a number of dynamic efficiency-based justifications for no challenge clauses, and illustrates that not every licensing practice that enhances patentee reward is justified from an innovation incentives perspective.

Finally, this article provides an overview of how the approach to these clauses in the major jurisdictions diverges from each other and from the analysis in this article. This divergence could be problematic because licensing is often done on a global basis, and divergent rules will result in a patchwork of licensing practices. This is doubly unfortunate, as this means that the applicable rules will instead be determined by the location of the licensees, even though most of the

goods produced by them are traded globally. This will result in an unnecessary distortion in the global market for many technological goods, an unfortunate outcome that is wholly avoidable by greater convergence in the regulation of patent licensing practices.

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 5 SPRING 2016 NUMBER 2

INTERVIEW

THE INDIGO BOOK: A MANUAL OF LEGAL CITATION

CHRISTOPHER SPRIGMAN & DARYL STEIGER

In this interview, NYU Law Professor Christopher Sprigman discusses his experience conceiving, drafting and producing the Indigo Book, a manual of legal citation that replicates the uniform system of citation expressed by Harvard Law Review's Bluebook. Formerly known as "Baby Blue," the Indigo Book is published by Public.Resource.Org under a creative commons license and may be freely modified or distributed without restriction under copyright law. Professor Sprigman details his vision for the Indigo Book, the legal copyright and trademark uncertainties the project faced, and his plans for the citation manual's future.

Christopher Sprigman has been a professor at NYU School of Law since 2013. Sprigman teaches intellectual property law, antitrust law, competition policy, and comparative constitutional law. His research focuses on how legal rules affect innovation and the deployment of new technologies. Sprigman's widely cited works have had an influence on important aspects of copyright law, and often belie the conventional wisdom about intellectual property rights. He was an appellate counsel from 1999 to 2001 in the Antitrust Division of the US Department of Justice, where US v. Microsoft was among his cases, and later was elected partner in the Washington, DC, office of King & Spalding. In 2003, Sprigman left private practice to become a residential fellow at Stanford Law School's Center for Internet and Society, and joined the faculty at the University of Virginia School of Law in 2005. Sprigman received his BA in history magna cum laude from the University of Pennsylvania in 1988, and a JD with honors from the University of Chicago Law School in 1993. He subsequently clerked for Judge Stephen Reinhardt of the US Court of Appeals for the Ninth Circuit and Justice Lourens H. W. Ackermann of the Constitutional Court of South Africa. Sprigman also taught at the University of the Witwatersrand's law school in Johannesburg.

Daryl Steiger is a student at NYU School of Law graduating in May 2016 and an Articles Editor for the NYU Journal of Intellectual Property and Entertainment Law. Daryl has focused her studies on intellectual property and entertainment law, and worked as a team lead on Professor Sprigman's project developing the Indigo Book, formerly known as Baby Blue. Daryl received her BA in women, gender and sexuality studies from Washington University in St. Louis in 2012, with a second major in psychology and a minor in film.

* * *

DS: Hi Professor Sprigman. Thank you for taking the time to do this. First, we were wondering if you would give us a brief overview of the Baby Blue project?

CS: Sure. It was back in late 2014 that I started discussing the possibility of this project with Carl Malamud, who is the head of something called Public Resource. Carl's mission in life is to bring law to the people that law regulates. He wants to make the laws of various states, as well as private codes, like building codes – all things that are often locked up in libraries and only available by subscription – he wants to make these things available online to people free of

charge. The feeling behind this is that both state laws and private codes that have been adapted as law regulate people's behavior, people are subject to their authority. So people should have the right to access these things and indeed courts have held that the law is not copyrightable and that edicts of law can be accessed free of copyright restrictions by the people who are subject to those edicts. So that's Carl's aim in life.

The Bluebook is something came to Carl's attention and he started thinking and, when he started talking to me, I also started thinking, that the Bluebook was kind of like a law, in that it ran a lot of people's lives. Associates at law firms, law students in law schools, academics in law schools; they all cite legal documents and they are obliged to do so according to the Bluebook's form. But the Bluebook isn't free – actually, it isn't free in two senses. First, it's expensive, and second, the system of citation that the Bluebook expresses is not open and is developed under the control of a small number of people at the Harvard Law Review. So we decided to think about whether we could do something about both of those problems. What we hit on was that the Bluebook is first of all a *system*, a system of legal citation. Systems are not in fact copyrightable. So we thought, well, let's see if we can express this same citation system in our own words. That is, we won't copy the expression but in fact re-express the system by explaining its rules differently. If we do that, then people can use our version of the Bluebook system to make legal documents that look exactly as if they were using the Bluebook. So we set out to try to do that.

The other thing we thought was, in the process, we could set the legal citation system free of copyright. One good thing that might come out of that is that people other than a tiny group of Harvard Law Review editors might work with this system, and ultimately simplify it and make it better. All interested members of the legal community would have a crack at modifying the system, improving it, updating it.

I agreed with Carl that I should take this on. I recruited a team of students here at the law school who worked with me, who carried the load on this and really did pull this system apart, understand it from the inside, and basically re-describe it in the way that I told you we would try to do. That's the project in a nutshell. The product is something that we first referred to as Baby Blue but now we're calling the Indigo Book. It is a Bluebook-compatible system of citation. Again – and this is crucial – people who use Indigo are using the Bluebook's system of citation. From the outside, no one will ever know they used the Indigo Book to produce their citations. For all intents and purposes, documents produced using the Indigo Book will look like a document that's been Bluebooked.

DS: Aside from Carl mentioning this to you, what personally motivated you to get so involved in this project?

CS: Part of it was just the copyright question. Could we take this system, which is not copyrightable – Section 102(b) of the Copyright Act says that copyright does not extend to systems – and could we understand it well enough to re-express it, to express it in different words? Now, sometimes rules are written the way they are and there's no way to really rewrite them. To make them understandable, you have to write the rule very similarly. That's interesting. That situation would lead you, as a copyright lawyer, to say that the expression of the rule has merged with the rule itself. You have to be able to express it the same way, because otherwise the rule would have a backdoor copyright, and rules don't have copyrights. So, I just thought as a copyright scholar this was a very interesting project.

And then I thought of people in jail. I thought of solo practitioners out there struggling to make ends meet. Then I thought, you know, the Bluebook's expensive and we should be able to bring a version of the system, Bluebook's system of citation, for free. And I also thought about all the people who had an interest in citations - practicing lawyers, academics, law students - who had no ability, if they're not on the Harvard Law Review, to actually say what the rules should be. I thought that's odd. By what authority does the Harvard Law Review get to proscribe rules for the rest of the lawyers in society? Not only lawyers, but pro se litigants: people filing briefs in criminal or civil cases who are not represented by lawyers. Who anointed the Harvard Law Review, right? So, this system is something I thought it something we can bring to people more efficiently, more cost effectively, and also in a way that they can work with it, to change it, to streamline it, to improve it. That was the reason why I decided to do this.

DS: So, as you just touched upon, one of the exciting things about the Indigo Book is that it's open source. How did you become involved with open source law generally? How did Carl know to approach you?

CS: So, I just want to be clear on what the licensing terms are for the Indigo Books. So the Indigo book has been released under something called a CC0 license, which is a creative commons public domain license. This means that the Indigo Book is in the public domain. People can copy it, they can distribute it, they can modify it, they can distribute their modifications. They can essentially do what they wish to do with it. The idea here again is that people can proliferate copies of this book cheaply and this will spread the book to other people. But also the idea is

that if people have thoughts on how to make a better citation system, how to reduce the complexity of legal citation – which would a good thing, legal citation is incredibly complex – they should and they are permitted to do that. They are permitted to make their modifications and to distribute them. And we're hoping over time that this will result in a lot more creative and helpful input into what legal citations should look like.

In terms of how I got involved in this, I'm a board member of creative commons, an organization I've supported for a long time and really believe in. I think creative commons brings to copyrighted works what the open source movement has brought to software, which is the ability for people to participate in determining the direction of some tools that all of us use, that some of us as lawyers use every day. Up until now, we have had no voice in what legal citation rules look like. Now we do. I think this is the genius of open source. It democratizes participation in the production and the maintenance of these tools – software tools, legal citation tools. That is something I think is just very good for all of us.

DS: So how did you approach drafting baby blue with you student team?

CS: The student team would meet with me on a regular basis and we'd think about, you know, what are the rules we need to express? We set the scope of the project. We understand what part of the Bluebook, that is the "Bluepages" that we were going to work from. We made a decision at some point that we wouldn't try to mess with the international part of the Bluebook. That part of the Bluebook was something that most lawyers in America don't really use. It's of limited utility, it lends a lot to the Bluebook's bulk, and if we try to restate that, then the thing we were producing, which eventually became the Indigo Book, would be similarly huge, and we wanted that not to happen.

So we set the scope of the project, then we split up assignments. We had different teams. We had teams named after animals. We had team Koala, Team Puma. Different teams would start working away on different pieces of this. And then the teams would start checking each other's work. Over time and many, many versions, we started working out a draft that was stable and then we started refining it. So this took a long time, this took more than a year of work, very hard work on behalf on a lot of students, with me organizing and then checking on all this work and going through and refining it.

Then we delivered it to Carl Malamud, who is the publisher, and he put this into HTML, and got this set up online and really made it look nice. This is the

product that we gave to people to comment on in a beta version. So that's how this actually happened.

DS: So, there were other people who were involved in the project besides Carl, yourself and the student members. What roles did they play and how did you get them involved in the project?

CS: We had a group of advisors. These were copyright experts – people like Pam Samuelson, for example, who was one of our advisors. She's a professor at Berkeley and really I think the figurative dean of the copyright professoriate. Pam gave us a fair amount of input, and others did as well, designed to make sure that when we made decisions about what our understanding of rules was like and how we express that understanding, that these decisions made sense. So we had a bunch of people who kept tabs on us and we could go to them with questions and they would give us input.

Also, we had Joe Gratz, who's a lawyer at Durie Tangri in San Francisco. Joe took a look at our work. He was one of the people advising us on our legal issues and then later on he represented us when we had discussions with the Harvard Law Review.

DS: So, if you feel comfortable discussing it, how did you approach the legal copyright concerns posed by the Indigo Book, what was formally known as Baby Blue, and how did you address these concerns when you were drafting it?

CS: So the concerns came in two varieties, mainly. One was copyright and the second was trademark. With respect to copyright, the overarching idea that I had from the beginning, which I think is right, was that the uniform system of citation, which is this legal citation system that the bluebook articulates, is a system. It's billed as such on the cover of the Bluebook, which says "uniform system of citation." Systems are simply not copyrightable. Section 102(b) of the Copyright Act rules them out of copyright. Systems you might be able to patent. The Bluebook is not a system you can patent. It's not a novel system. It's been on sale forever. It would fail if they tried to patent it.

So it's not a patentable system and systems are not copyrightable. If that's the case, then the rules, the structure of rules that makes up the uniform system of citation is basically free for all of us to use. The copyright question isn't really about the rules. The copyright question is really about how you express them. Really what the project was about was understanding what the rules are and trying to re-express them in language that made the rules clear without being identical to the language in the Bluebook. Now, when we needed to express a rule in a certain

way, we did. And we would depend on the merger doctrine to shelter us. It gives us the ability to do that because if you can't have a copyright on a rule and there's only one way to express the rule, you got to be able to express it. So where we needed to rely on the merger doctrine, we did.

We also replaced all of the examples. The Bluebook is full of examples. We tried to make them better where we could. We picked different examples and we tried to make the examples a little bit better, sometimes a little bit funny. We added a bunch of material trying to clue people into the real meaning of the rules – what we first called Baby Blue clues and then we ended up calling the Indigo Inklings. So these were little hints we dropped into the book at different points to help people understand how the rule works and how to use the rule. If you read the Indigo Book, it just reads very different than the Bluebook. It's more succinct. I think it's more entertaining. I think it's sleeker and more efficient and just a better citation manual for what it does.

The second issue is trademark. Now, I wanted this book to be called Baby Blue for a reason. The Bluebook is Harvard Law Review Association's version of the uniform system of citation. Most people, if you talk to them, they don't know what the uniform system of citation is, they only know the Bluebook. It was important to me that we find a way to signal that what we were doing was compatible with the Bluebook, that it was in fact an expression of the same uniform system of citation that the Bluebook is an expression of. Therefore, if you used our manual, your documents would look just like you were using the Bluebook. No one would be able to tell. I thought this was important because, you know, lawyers are very conservative. The Bluebook is the standard. We wanted to recreate the standard. We wanted to give people a choice in how they used the standard – the uniform system of citation. They could use it by using the Bluebook or they could employ that system by using what we called Baby Blue.

So I thought Baby Blue was a good name in that it communicated to people that this book was "blue." That you could use this book and your documents would be consistent with the Bluebook's uniform system of citation. I thought of an example. If you think of artificial sweetener packets, "Equal" is the blue one, "Splenda" is the yellow one, and "Sweet 'N Low" is the pink one. People don't know the substance in that blue packet is aspartame, right? People don't remember aspartame. They remember either the brand name, "Equal," or they remember the color of the package, blue. So competitor generic aspartame packets are often blue. The reason is because people are looking for the blue sweetener. In the same way that competition in artificial sweeteners depends on the use of a color that people have come to associate with a particular product, I think competition, to my mind,

in the market for legal citation systems requires a reference to this color, which has become the color of a particular form of legal citation. That's why I thought "Baby Blue" was a good name.

Now, I didn't think Baby Blue would be confusing to people. I thought people would understand that Baby Blue was a very different thing than the Bluebook. We had a disclaimer on the cover that said that this is not sponsored by or affiliated with the Harvard Law Review or the Bluebook. But the Harvard Law Review people threatened to sue. They were adamant that we not use the word "blue" in our title. Eventually, we decided that we would rather this book come out and have people start using it than spend a couple of years fighting this out with Harvard. Although I'm confident we would have beaten them, I didn't want to spend a couple years with Harvard fighting it out.

So we changed the name to the "Indigo Book," which I don't think, under anybody's account of trademark law, would be a trademark violation. I can't imagine people being confused between the Indigo Book and the Bluebook. Now indigo of course is the plant that is used to make a form of blue dye. It is a dark, kind of purplish-blue color. We are trying to signal to people that this is blue. I'm not sure if it's as effective a signal as "Baby Blue" was. We'll see. But I do need people to understand that this book is compatible with the Bluebook. That's the message that we're trying to send.

DS: Going back to your motivation, were you inspired in any way by Richard Posner's article, "Bluebook Blues"?

CS: Yeah, so, Richard Posner really hates the Bluebook. Recently he actually said that he believes the Bluebook should be burned, which I think is fairly extreme. I'm generally anti-book burning. So, you know, the University of Chicago has some experience with this. When I was a student there – I went to law school there – there was something that Chicago published called the Maroonbook. The Maroonbook was an elegant little book of legal citation. Very skinny. Not really too many rules. There were just the right number of rules and then there was a standard, which was, you know, when a rule doesn't cover a situation, try to make yourself understood. Try to construct a citation that will allow people with some reasonable effort to locate the source, which I think is a great way to think about citation.

Like a lot of things University of Chicago does, it was very terrific, and I think very thoughtful, and it failed. It gained basically no traction. And the reason again I think is the deep conservatism of lawyers. Generations of lawyers have been educated to use the Bluebook. That's the resource they know how to use, they

don't want to switch. Even though I think the Maroonbook is a much better resource, it didn't gain any ground. So this really did inform us. Judge Posner has been agitating for a simpler form of legal citation. I very much agree with him but I think this is a goal that's best approached in two steps.

The first step is to set the Bluebook free of copyright. That's what I think we've done with the Indigo Book. This is an implementation of the uniform system of citation that anyone can work with. By "work with" I mean do what Judge Posner thinks should be done: simplify dramatically. There's no reason that there should be a rule articulated for every circumstance. You could probably cover the majority of citation questions within a dozen rules. Then everything else could be left to a standard, which is to construct a citation that would allow a reasonably smart person to find the resource reasonably well. We can get there. Before we do that, we need to get to step one, which is to break the monopoly that Bluebook has had over the uniform system of citation, which is the underlying citation system, which is not copyrightable. There has to be some other way that people can access this system – a way they can, not simply access, but that they can work with, that they can change, that they can streamline, that they can improve, that they can expand. This is the strategy.

In terms of goals, I'm completely with Judge Posner. In terms of strategy, I think we would differ a bit.

DS: So now that the Indigo Book is out, and it's out to the public, what feedback have you received from the legal community and what impact do you hope it will have on the legal profession as a whole?

CS: I mean, people like it. They like it a lot. We've gotten a lot of feedback from people who are using it and enjoying using it. I think now, truthfully, the work has just started. We have the resource. It's available to most people. A whole bunch of work needs to be done. Whether I'm the one to do it is unclear. This is a job probably for a lot of people with a lot of different talents, bringing a lot of different ideas to the task.

So, it's important that the next generation of law students knows that there is a choice available to them. They can use the Bluebook, they can use the Indigo Book, and they'll get their job done either way. It's important that lawyers out in the marketplace know this too. There's going to be a lot of evangelism that goes into making sure that they do know this.

Equally important is the task of attracting developers to the Indigo Book; people who are going to think about how to make citation better. And how to

revise this book and circulate the revisions in a way that makes the process of determining our legal citation rules more democratic. This work all needs to be done and we've just gotten to the first step. We're going to be busy.

DS: So, going beyond that step, what's next for the Indigo Book? Do you have any plans to support and update it?

CS: We're trying to figure that out right now. Like any open source project, it's not really about us. It's about all the people who want to be involved. So, part of the challenge moving forward would be to attract people to the task of developing extensions, changes to the Indigo book. That's something we can do using open source tools.

So, it's to be determined. There are a bunch of different ways to think about the development of this resource and we're just trying to figure out what's the right way to do it at this point. We don't have yet a firm plan, but we're working on it. A lot of that thinking will be done this summer.

DS: I'm excited to hear what goes on then. So, where else could the legal profession benefit from an open source approach?

CS: Well, case law. You know a lot of the legal profession accesses case law through Westlaw and Lexis. They pay for the privilege. There's no reason that these government works, which we pay for the production of through our tax dollars, should be locked up in a Lexis / Westlaw duopoly. A lot of case law is now available on the web, but not nearly all. Making case law more widely available will open up the law to people who don't have access to it through the expensive, traditional means. People are working on this, that's a project that's ongoing, and there's a lot of hands on that, so I think there will be progress over time.

DS: What about court filings? Do you think that would be an area that also needs it?

CS: Yes. So, PACER is the service that the federal courts use for access to court filings. I know a lot of people who use PACER to gain access to court filings and they get their credit card bills at the end of the month. It's shocking how quickly your credit card bills mount when using PACER. Again, I understand that the federal court system needs to recover its costs – the cost of providing court filings to the public – but PACER just seems like an extraordinarily complex and expensive way to do that.

DS: In all, what lessons do think the Indigo Book poses for copyright reform generally?

CS: Yeah, so I haven't really thought about that. I don't think of this as a copyright reform project. I think of this more of a project that attempts to make some infrastructure that's very important to the public available to the public. Not simply available to use, but available to change, to work with, to govern. So I don't have a broader copyright reform idea in terms of how the Indigo Books relates to this topic. It's really a project that has to do with a specific piece of infrastructure that is important to lawyers.

DS: Finally, before we close, how did this project or open source law in general relate to your professional or academic work?

CS: I write a lot about areas in which there isn't a lot of copyright, but there's a lot of innovation, so the fashion industry is a great example of that, open source software is another. I think, in our legal citation system, there hasn't been a lot of innovation. There's been a lot of growth in the number of rules, but the number of rules is not how you judge a legal citation system. Its flexibility, its suitability to the task, its accessibility by lawyers and by law students—that's how you judge it. I think we could have a lot of innovation in all those areas. I actually think that an open source approach is more likely to bring us that innovation than continued control by one unrepresentative and small group of elite law students.