

SCOPE[†]

MARK A. LEMLEY* & MARK P. MCKENNA**

ABSTRACT

Virtually every significant legal doctrine in IP is either about whether the plaintiff has a valid IP right that the law will recognize (validity); whether the defendant's conduct violates that right (infringement); or whether the defendant is somehow privileged to violate that right (defenses). IP regimes tend to separate doctrines in these three legal categories relatively strictly. They apply different burdens of proof and persuasion to infringement and validity. In many cases they ask different actors to decide one doctrine but not the other. And even where none of that is true, the nature of IP law is to categorize an argument in order to apply the proper rules for that argument.

The result of this separation is that parties treat IP rights "like a nose of wax, which may be turned and twisted in any direction." When infringement is at issue, IP owners tout the breadth of their rights, while accused infringers seek to cabin them within narrow bounds. When it comes to validity, however, the parties reverse their

† © 2016 Mark A. Lemley & Mark P. McKenna.

* William H. Neukom Professor, Stanford Law School; Partner, Durie Tangri LLP.

** Associate Dean for Faculty Development, Professor of Law, and Notre Dame Presidential Fellow, Notre Dame Law School.

Thanks to Barton Beebe, Michael Burstein, Julie Cohen, Graeme Dinwoodie, Kristy Downing, Rochelle Dreyfuss, Brett Frischmann, Jeanne Fromer, Jim Gibson, Andrew Gilden, Paul Goldstein, Rose Hagan, Paul Heald, Laura Heymann, Tim Holbrook, Jeff Lefstin, Yvette Liebesman, Mike Madison, Tyler Ochoa, Sean O'Connor, Lisa Ramsey, Michael Risch, Zahr Said, Pam Samuelson, Dave Schwartz, Lea Shaver, Jessica Silbey, Christopher Sprigman, Eva Subotnik, Kathleen Sullivan, Rebecca Tushnet, Melissa Wasserman, Steve Yelderman, and participants in the Works in Progress-IP Conference, the University of Illinois IP Colloquium, the NYU Innovation Policy Colloquium, the Association of American Law Schools Annual Meeting, and workshops at Florida State, Indiana University McKinney School of Law, and Saint Louis University for helpful discussions of these issues.

positions, with IP owners emphasizing the narrowness of their rights in order to avoid having those rights held invalid and accused infringers arguing the reverse. Because of the separation between validity, infringement, and defenses, it is often possible for a party to successfully argue that an IP right means one thing in one context and something very different in another. And courts will not necessarily detect the problem because they are thinking of only the precise legal issue before them. The result is a number of IP doctrines that simply make no sense to an outsider.

The culprit is simple, but fundamental: IP regimes largely lack an integrated procedure for deciding the proper extent of an IP right. Without some way of assessing the breadth of an IP right that considers validity, infringement, and defenses together, courts will always be prone to make mistakes in applying any one of the doctrines separately. In this Article, we suggest that IP regimes need a process for determining the scope of an IP right.

Scope is not merely validity, and it is not merely infringement. Rather, it refers to the range of things the IP right lawfully protects against competition. Only by evaluating scope in a single, integrated fashion can courts avoid the nose of wax problem that has grown endemic in IP law. Scope is, quite simply, the fundamental question that underlies everything else in IP law, but which courts rarely think about expressly.

TABLE OF CONTENTS

INTRODUCTION	2200
I. IP LITIGATION: A PROCESS DIVIDED	2204
A. <i>The Scope of IP Rights</i>	2204
B. <i>Why Scope Matters Now</i>	2214
II. THE PROBLEMS WITH SEPARATION	2219
A. <i>Gaps in the Fabric of IP Scope</i>	2219
B. <i>Gaps Exist in the Fabric of IP for a Number of Reasons</i>	2220
1. <i>Different Decisionmakers</i>	2220
2. <i>Different Burdens of Proof</i>	2222
3. <i>A Reluctance to Invalidate IP Rights</i>	2224
4. <i>Strategic Behavior</i>	2225
C. <i>IP Doctrines and the Nose of Wax</i>	2225
1. <i>Copyright</i>	2226
2. <i>Patent</i>	2240
3. <i>Trademark</i>	2243
4. <i>Design Patent</i>	2259
III. TOWARDS AN INTEGRATED DOCTRINE OF IP SCOPE	2266
A. <i>The Need for an Integrated Scope Determination</i>	2266
B. <i>Markman as a Scope Proceeding</i>	2268
C. <i>Adapting the Scope Proceeding for Other IP Rights</i>	2271
D. <i>Alternatives to a Scope Proceeding</i>	2282
CONCLUSION	2285

INTRODUCTION

Intellectual property (IP) law doctrines fall into three basic categories: validity, infringement, and defenses. Virtually every significant legal doctrine in IP is either about whether the plaintiff has a valid IP right that the law will recognize (validity); whether what the defendant did violates that right (infringement); or whether the defendant is somehow privileged to violate that right (defenses).¹

IP regimes tend to separate doctrines in these three legal categories relatively strictly. They apply different burdens of proof and persuasion to infringement and validity. In many cases they ask different actors to decide one doctrine but not the other. The U.S. Patent and Trademark Office (PTO), for example, decides questions of patent and trademark validity but not questions of infringement. Even in court, resolution of one issue is often allocated to a judge while the jury decides a different issue. And even if none of that is true in a given case, the nature of IP law is to categorize an argument in order to apply the proper rules for that argument.

The result of this separation is that parties treat IP rights “like a nose of wax which may be turned and twisted in any direction.”² When infringement is at issue, IP owners tout the breadth of their rights, while accused infringers seek to cabin them within narrow bounds. When it comes to validity, however, the parties reverse their positions: IP owners emphasize the narrowness of their rights in order to avoid having those rights held invalid, and accused infringers argue the reverse.³

Because of the separation between validity, infringement, and defenses, a party may often successfully argue that an IP right means one thing in one context and something very different in

1. If the IP owner prevails, there are also issues about the remedy awarded. We do not consider remedial issues in this Article. Nor do we consider trade secrets or the right of publicity, even though both are arguably IP rights and implicate the questions we consider. We'd say they are outside the scope of our Article, but that would be a bit too meta even for us.

2. *White v. Dunbar*, 119 U.S. 47, 51 (1886).

3. See, e.g., Oren Bracha, *The Ideology of Authorship Revisited: Authors, Markets, and Liberal Values in Early American Copyright*, 118 *YALE L.J.* 186, 267-68 (2008) (“[Lawyers] can argue doctrinal questions of originality in terms that assume a cumulative process of creation and switch without blinking to arguing questions of infringement in terms that assume a sharp distinction between an original and a derivative.”).

another. And decisionmakers won't necessarily detect the problem because they are thinking of only the precise legal issue before them.

The result is a number of IP doctrines that simply make no sense to an outsider. In patent law, for instance, it is accepted law that there is no "practicing the prior art" defense.⁴ In other words, one can be held liable for doing precisely what others had legally done before, even though a patent isn't supposed to cover things people have already done. In design patent law, one can be held liable for making a design that an "ordinary observer" would find too similar to a patented design, even though the things that make the two look similar—say, the roundness of the wheels on a car—are not things the patentee is entitled to own. In copyright, once a court has concluded that the defendant has actually copied from the plaintiff, her song may be deemed infringing because of its similarity to the plaintiff's, even if the similarity is overwhelmingly attributable to unprotectable standard components of the genre. And in trademark, a party can be deemed infringing because its products look too similar to the plaintiff's and therefore make confusion likely, even if that confusion is attributable to non-source-designating features of the design.

Because they are sometimes sensitive to this problem, courts have created various stopgap doctrines to try to deal with these holes in IP law. In copyright law, for instance, expression cannot be protected if there are few alternative means of expression, such that one author's way of putting something has "merged" with the idea she is expressing.⁵ Trademark law has tried to prevent overreaching in product configuration protection by making it hard to have a valid product configuration trade dress at all.⁶ Patent disclosure doctrines (written description and enablement) try to solve this problem with patent claiming.⁷ But even these efforts to patch the system are doomed to fail at determining the proper scope of an IP right for the simple reason that they are themselves either validity

4. *See infra* Part II.C.2.

5. *See, e.g.,* *Morrissey v. Procter & Gamble Co.*, 379 F.2d 675, 678 (1st Cir. 1967).

6. *See, e.g.,* *Wal-Mart Stores, Inc. v. Samara Bros., Inc.*, 529 U.S. 205, 216 (2000).

7. For a discussion of claiming, see Timothy R. Holbrook, *Patents, Presumptions, and Public Notice*, 86 IND. L.J. 779, 809 (2011).

or infringement doctrines and are therefore subject to the same nose of wax problem as different decisionmakers tackle different pieces of the issue at different times.

The culprit is simple, but fundamental: IP regimes largely lack an integrated procedure for deciding the proper extent of an IP right. The proper scope of an IP right is not a matter of natural right or immutable definition. Rather, it is a function of the purposes of the IP regime. But without some way of assessing the breadth of an IP right that considers validity, infringement, and defenses together, courts will always be prone to make mistakes in applying any one of the doctrines. In general, these mistakes tend to redound to the benefit of IP owners, as courts are more reluctant to declare an IP right invalid altogether even when the defendant should not be liable for infringement. But sometimes the mistakes run the other way, as when the plaintiff's improper attempt to expand the scope of her legitimate but narrow right leads a court to hold that right altogether invalid.

In this Article, we suggest that IP regimes need a process for determining the *scope* of an IP right. Scope is not merely validity, and it is not merely infringement. Rather, it refers to the range of things the IP right lawfully protects against competition.⁸ IP rights that claim too broad a scope tend to be invalid, either because they tread on the rights of those who came before or because they cover things that the law has decided no one should own. IP rights with narrower scope are valid, but the narrowness of that scope should be reflected in the determination of what actions do and do not infringe that right. And whatever the doctrinal label, we should not allow an IP owner to capture something that is not within the

8. Some academic works think about the scope of an IP right in an integrated way, although usually only with one IP regime in mind and without our focus on the gaps between validity, infringement, and defenses. *See, e.g.*, Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1675 (2003) (evaluating patent policy levers in part on the basis of how they affect the scope of rights in different industries); Bernard Chao, *The Infringement Continuum*, 35 CARDOZO L. REV. 1359, 1359 (2014) ("For many years, patent law has struggled with the issue of permissible claim scope."); Mark A. Lemley, *The Changing Meaning of Patent Claim Terms*, 104 MICH. L. REV. 101, 119-21 (2005) (discussing the scope of patent claims in an integrated way); Zahr K. Said, *Reforming Copyright Interpretation*, 28 HARV. J.L. & TECH. 469, 516 (2015) (arguing for a judicial determination of scope in copyright proceedings). Chao goes on to suggest calibrating remedies as a way to modulate claim scope. Chao, *supra*, at 1404-07.

legitimate scope of her right. Nor should it follow from the fact that some uses are outside the lawful scope of an IP owner's right that the IP right itself is invalid and cannot be asserted against anyone. Only by evaluating scope in a single, integrated proceeding can courts avoid the nose of wax problem that has grown endemic in IP law. Scope is, quite simply, the fundamental question that underlies everything else in IP law, but which courts rarely think about expressly.

One IP regime—utility patent law—has started in the last two decades to think about scope in a more systematic way through the process of claim construction. Patent courts hold a pretrial *Markman* hearing to determine what the patent does and does not cover; they use that determination to inform both validity and infringement.⁹ Claim construction has its share of problems,¹⁰ and patent law has not entirely avoided the nose of wax problem, but we think patent law is on the right track in trying to resolve these questions in an integrated fashion rather than allowing inconsistent assessments of scope in separate validity and infringement proceedings.¹¹ Other IP regimes can look to a modified form of the *Markman* model as they design their own doctrines of scope, learning from its shortcomings while trying to modulate the breadth of permissible IP rights.

In Part I, we discuss the current divided structure of IP regimes, explaining how validity, infringement, and defenses are often evaluated at different times by different parties applying different standards. We also explain why this separation has taken on increased importance in recent years. In Part II, we discuss the problematic consequences of that separation for each of the major IP regimes. Finally, in Part III we propose that IP adopt a unified scope proceeding, drawing inspiration—but also some caution—from patent law's *Markman* experiment.

9. See *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 384-91 (1996).

10. For a discussion, see Dan L. Burk & Mark A. Lemley, *Fence Posts or Sign Posts? Rethinking Patent Claim Construction*, 157 U. PA. L. REV. 1743, 1744 (2009).

11. See Lemley, *supra* note 8, at 122.

I. IP LITIGATION: A PROCESS DIVIDED

A. *The Scope of IP Rights*

IP cases have a familiar tripartite structure. The first question is whether the plaintiff owns any valid rights. If (and only if) the plaintiff does have valid rights, then the second question is whether the defendant's conduct infringes those rights. Third is the question of whether the defendant might have some defense or special exclusion that exempts from liability its otherwise infringing conduct. More particular doctrines are then conceived of as validity or infringement issues, or as matters of defense. The allocation into one of these buckets matters because the consequences of treating some doctrine as a validity rather than an infringement doctrine are often different. Validity doctrines tend to have all-or-nothing effect, so a finding against the claimant on one of these doctrines typically leads to the conclusion that the claimant has no rights at all. Infringement doctrines, on the other hand, tend to focus on the conduct of a particular defendant, so a finding on one of these doctrines typically applies only as against that defendant. Validity and infringement doctrines also may be decided by different decisionmakers.

This kind of rigid separation between validity, infringement, and defenses often doesn't work, however, because most of the important questions in modern IP cases are really questions of the *scope* of a party's rights, and scope has a variety of different dimensions that don't fit neatly into these doctrinal categories. Indeed, though they tend not to get headline billing, questions of scope are pervasive throughout IP cases. And the correct scope of an IP right is not found in nature; it is determined by a series of policy judgments.

In each area, courts first must determine the thing to which rights might attach: What is the invention, the design, the work, or the mark?¹² This is, in the first instance, a question of what a party claims to own. But it's not that simple. For one thing, the IP systems differ from each other in their claiming requirements—in

12. Cf. Michael J. Madison, *Law as Design: Objects, Concepts, and Digital Things*, 56 CASE W. RES. L. REV. 381, 383 (2005) (discouraging the focus on "rights in things" and encouraging the examination of "the origins of things").

terms of timing, claiming methodology, and specificity.¹³ Patent owners arrive at court with a document that (however unclearly) spells out their claims in writing or, in the case of design patents, in pictures. Copyright owners allege infringement of some work of authorship that can serve as a starting point, but they have not articulated specifically which aspects of that work warrant protection. Trademark owners sometimes assert rights in registered marks, for which the registration depicts the mark and identifies the goods and services in connection with which the mark is used. But plenty of trademark claimants assert rights in unregistered marks, the boundaries of which may be unclear. And uncertainty remains even in cases involving registered marks. For one thing, courts typically focus on the nature of the plaintiff's use rather than the form of the mark or goods and services identified in the registration.¹⁴ And the identified goods and services would only be the beginning of the inquiry even if courts did focus on the registration because the defendant need not use its mark for identical goods; its use need only be similar enough that confusion is likely (in consideration of other factors).

This first kind of scope assessment aims to define the invention, work, or mark at issue, primarily so that a court can determine whether the claimed property is subject to valid rights at all. One significant part of the validity determination in every area of IP is a comparison of the claimed IP with what came before it. Almost all IP regimes premise protection on some form of novelty—the invention or creation of a new thing the world has not seen before.¹⁵ But truly new creations are rare things. Almost all IP owners add their contributions onto a base of prior knowledge.¹⁶ And, for various

13. See Jeanne C. Fromer, *Claiming Intellectual Property*, 76 U. CHI. L. REV. 719, 719 (2009).

14. See Rebecca Tushnet, *Registering Disagreement: Registration in Modern American Trademark Law*, 130 HARV. L. REV. (forthcoming 2016) (manuscript at 40-46), <http://ssrn.com/abstract=2735013> [<https://perma.cc/LP5M-J9BA>].

15. Even trademark law, which does not require novelty per se, typically awards rights to the party that first uses a mark in connection with particular goods or services. For example, Apple Inc. did not invent the term “apple,” but it was the first to use it to indicate the source of computers.

16. Copyright, of course, does not require novelty in the sense that the work must be different from what came before it. Indeed, as a matter of black letter doctrine, copyright attaches to any work independently created (not copied from others) with a modicum of

policy reasons, IP regimes simply do not count some features, no matter how different they might be from what came before them. Patent law will not protect laws of nature or abstract ideas, even if they are novel and nonobvious;¹⁷ copyright law will not protect ideas or processes no matter how original they are;¹⁸ trademark law will not protect functional features even if they are source-indicating.¹⁹ Thus, one important purpose of asking “what is it?” is to be able to test whether valid rights attach in relation to a unit of the claimed scope, and in particular which features of the claimed property are protectable.²⁰

Importantly, all IP regimes recognize that a party can own some valid rights in a work despite the fact that many features of the work are not protectable. One can, for example, patent an invention that incorporates many elements of prior inventions, as long as the newly claimed invention adds some novel and nonobvious new element.²¹ Similarly, one can own a valid copyright on expressions of facts or on photographs of well-known places,²² just as one can claim

creativity. But that rule is more theoretical than practically applicable—courts almost universally treat “striking similarity” between works as strong evidence of copying. And more to our point, the scope of copyright protection in a work that is identical to other preexisting works would be vanishingly small. The plaintiff would have to prove that the defendant’s work was copied *from the plaintiff* rather than from the preexisting work.

17. *Alice Corp. Pty. Ltd. v. CLS Bank Int’l*, 134 S. Ct. 2347, 2354 (2014).

18. 17 U.S.C. § 102(b) (2012).

19. *Qualitex Co. v. Jacobson Prods. Co.*, 514 U.S. 159, 164-65 (1995).

20. *See* Madison, *supra* note 12, at 397-98; Michael J. Madison, *IP Things as Boundary Objects: The Case of the Copyrighted Work* 33-34 (Univ. of Pittsburgh Legal Studies Research Paper Series, Working Paper No. 2013-12, 2013), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2256255 [https://perma.cc/RB36-CZ5X].

21. *Cf.* *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418-19 (2007) (“Although common sense directs one to look with care at a patent application that claims as innovation the combination of two known devices according to their established functions, it can be important to identify a reason that would have prompted a person of ordinary skill in the relevant field to combine the elements in the way the claimed new invention does. This is so because inventions in most, if not all, instances rely upon building blocks long since uncovered, and claimed discoveries almost of necessity will be combinations of what, in some sense, is already known.”).

22. *See* *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 348 (1991) (“Factual compilations, on the other hand, may possess the requisite originality. The compilation author typically chooses which facts to include, in what order to place them, and how to arrange the collected data so that they may be used effectively by readers. These choices as to selection and arrangement, so long as they are made independently by the compiler and entail a minimal degree of creativity, are sufficiently original that Congress may protect such compilations through the copyright laws.”).

trademark rights in phrases with descriptive words or product packaging with some standard features.²³ In patent law we refer to the features that make an invention patentable collectively as the “point of novelty.”²⁴ Copyright and trademark law lack a similar term, but they have the same concept. Those features that are new and are not excluded from protection for policy reasons can be the basis of protection for a work as a whole, but that which was taken from the prior art or otherwise excluded from protection cannot be protected separately.

These determinations, which together define the nature of a party’s entitlement, are in some ways analogous to the ones that frame disputes about real property. In order to determine whether a party has violated the owner’s rights in a piece of property, a court must first determine the metes and bounds of the property claimed and the interest the claimant has in that property (that is, was the property taken in fee simple, or is the interest a future interest, a lease, etc.?). Similarly, to determine whether a party has violated an IP owner’s rights, a court must first determine what constitutes the invention, the work, or the mark. But unlike real property, no matter what the claiming rules, delineating the boundaries of the IP right requires *interpretation*, and that interpretation is necessarily done in the shadow of legal rules. That leads us to the second type of scope inquiry, which focuses on the nature of a party’s rights in the property.

Once we know what the IP owner has in fact contributed to the world, we can ask what acts violate rights in that property. This inquiry also has embedded within it several questions related to scope. Every IP system determines infringement by reference to both the acts that cannot be undertaken in relation to the subject of IP rights and the level of the similarity between the defendant’s invention, work, or mark and that of the plaintiff.²⁵ Patent, for

23. See *Antioch Co. v. W. Trimming Corp.*, 347 F.3d 150, 158 (6th Cir. 2003) (“[I]n order to receive trade dress protection for the overall combination of functional features, those features must be configured in an arbitrary, fanciful, or distinctive way.”); 2 J. THOMAS MCCARTHY, *MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION* § 11:26 (4th ed. 2015) (“The combination of two or more admittedly descriptive elements as a composite mark may result in a composite which is nondescriptive.”).

24. See Mark A. Lemley, *Point of Novelty*, 105 NW. U. L. REV. 1253, 1254 (2011).

25. Jeanne Fromer and Mark Lemley speak of IP regimes as requiring some combination

example, prohibits others from making, using, offering to sell, or selling the patented invention within the United States, or importing the invention into the United States.²⁶ That formulation requires both a determination that the defendant is engaged in certain acts—the making, using, selling, and so on—and a determination that the defendant is engaging in those prohibited acts with “the patented invention,” or, in the case of design patent, the “patented design.” Thus, a party that makes the patented invention outside the United States (and does not import it) does not infringe. Nor does a party that sells within the United States an invention that does not read on each element of the claimed invention.

Similarly, copyright requires a determination that the defendant has reproduced, distributed to the public, publicly displayed, or publicly performed, “the copyrighted work,” or that it has created a derivative of that copyrighted work.²⁷ A party that privately performs the copyrighted work does not infringe. Nor does a party that publicly performs something that is not the copyrighted work. Trademark law requires that the defendant make commercial use of the plaintiff’s mark in a way that is likely to cause confusion or to dilute a famous mark.²⁸ A party that does not make commercial use does not infringe, even if whatever use it made was likely to cause confusion. Likewise a party that makes commercial use of a mark does not infringe if that mark is too dissimilar or is used in a context that makes confusion and dilution unlikely.

Together the prohibited acts and the requisite similarity determine the legal scope of a party’s rights.²⁹ But of course that legal scope is inextricably intertwined with the delineation of the entitlement we previously described. For when we say that the defendant must publicly perform “the copyrighted work,” we take for granted that we have already been able to define “the copyrighted work” so that we can anchor our analysis of similarity in relation to that

of “technical similarity” and “market substitution.” Jeanne C. Fromer & Mark A. Lemley, *The Audience in Intellectual Property Infringement*, 112 MICH. L. REV. 1251, 1251 (2014).

26. See 35 U.S.C. § 271(a) (2012).

27. 17 U.S.C. § 106 (2012).

28. See 15 U.S.C. §§ 1114, 1125 (2012).

29. To be more precise, they determine some dimensions of legal scope. A variety of other considerations come into play to fully delineate legal scope—the duration of rights and their geographic scope, for example.

work. Further, we generally mean that the defendant must have taken the thing that gives the IP right its validity—the point of novelty, in patent terms. The marks “Shake Shack” and “Joe’s Crab Shack” might share the word “shack,” but that fact alone shouldn’t cause the newcomer to be guilty of trademark infringement, for the simple reason that the word “shack” alone is not what justifies protecting the trademark. Nor does the fact that our title includes the word “scope” give a mouthwash company the right to sue us for trademark infringement.³⁰ Similarly, your novel might share with our Article the word “the,” but copying that word from us doesn’t make you a copyright infringer, because the originality of our Article does not result from our use of the word “the” by itself.

Determination of the scope of a party’s rights therefore necessarily depends on a definition of the relevant IP. More particularly, evaluating whether the defendant’s use comes within the legal scope of the plaintiff’s rights requires identification of the protectable elements of those rights, because all of the IP regimes require, at least in theory, not just similarity between the defendant’s and plaintiff’s works, but similarity with respect to the protectable elements.

Our primary focus here is on the question of how courts manage scope in this respect—how they determine the protectable elements of a work and ensure that infringement analysis in fact focuses on those protectable elements. But of course we recognize that this is only one part of the scope story.

Legal scope is also determined in part by various defenses. For one thing, some doctrines, even though formally classified as defenses, really go to the question of validity. Certain issues, for example, are only formally designated defenses because the IP at issue is registered and therefore benefits from a presumption of validity, putting the burden on the defendant with respect to that issue.

30. These examples sound ridiculous. But it is surprisingly common for trademark owners to assert claims against marks that resemble their own only in relation to obviously unprotectable elements. *See, e.g.,* M.Z. Berger & Co. v. Swatch AG, 787 F.3d 1368 (Fed. Cir. 2015) (considering allegation that “iWatch” for watches was too similar to “Swatch”); Complaint, *Ky. Mist Moonshine v. Univ. of Ky.*, No. 5:15-cv-00328-JMH (E.D. Ky. Nov. 2, 2015) (challenging allegation that “Kentucky Mist Moonshine” for whisky was too similar to the University of Kentucky’s registration of “Kentucky” for educational and sports services); Complaint, *Star of Tx. Fair & Rodeo v. El Rodeo Inc.*, No. 1:15-cv-00992 (S.D. Tex. Nov. 2, 2015) (alleging that “El Rodeo” for a restaurant was too similar to “Rodeo Austin” for a rodeo).

Distinctiveness for registered trademarks is presumed, so a defendant that claimed the asserted mark lacked distinctiveness would have to raise that as an affirmative defense and would bear the burden of proof on that issue.³¹ Invalidity in a patent case similarly must be raised by the defendant as a defense.³² Because these kinds of defenses really relate to validity, they tend to be all-or-nothing, meaning that these defenses generally go to the existence of rights at all, rather than simply excusing the particular defendant's conduct.³³ Classifying these doctrines as defenses matters, however, because it affects the allocation of the burden of proof and how heavy that burden will be.³⁴ Invalidity in patent cases, for example, must be proved by clear and convincing evidence.³⁵ Still, defenses of this type require decisions about whether the IP rights should exist at all.

Other defenses are more like excuses in that they don't deny the existence of valid rights but offer some reason why those rights do not reach this particular defendant under the circumstances. In copyright law, the statutory exemption for using protected works in face-to-face classroom teaching does not alter the copyright owner's rights vis-à-vis any other use, even other educational uses, which usually must be adjudicated under the less specific standards of fair use.³⁶

31. See 15 U.S.C. § 1115(a).

32. See 35 U.S.C. § 282(a) (2012).

33. There are some exceptions here, particularly in trademark law, which allows unfair competition remedies in very limited situations even when a mark has been found generic. See, e.g., *Kellogg Co. v. Nat'l Biscuit Co.*, 305 U.S. 111, 116, 122-23 (1938); *Blinded Veterans Ass'n v. Blinded Am. Veterans Found.*, 872 F.2d 1035, 1043 (D.C. Cir. 1989); *King-Seeley Thermos Co. v. Aladdin Indus.*, 321 F.2d 577, 581 (2d Cir. 1963). The injunctions in these cases are quite narrow, and they specifically cannot prevent competitors from using the generic terms altogether. See *Blinded Veterans*, 872 F.2d at 1043 ("[T]he subsequent competitor cannot be prevented from using the generic term to denote itself or its product, but it may be enjoined from passing itself or its product off as the first organization or its product. Thus, a court may require the competitor to take whatever steps are necessary to distinguish itself or its product from the first organization or its product.").

34. Calling something a defense may also affect who decides—though as we note below, the allocation of decisionmaking responsibility is surprisingly unclear in some of the contexts with which we're concerned.

35. See *Microsoft Corp. v. i4i Ltd. P'ship*, 131 S. Ct. 2238, 2242 (2011).

36. See 17 U.S.C. § 110 (2012). And, of course, despite its more general standards, fair use itself focuses on the legitimacy of the specific defendant's use rather than denying the existence of any rights.

These different types of defenses exist in varying proportion in different IP regimes. Utility patent law has many of the former type of defenses because patent requires defendants to prove invalidity and unenforceability, but it has virtually none of the latter type, save for the extremely narrow experimental use defense.³⁷ Design patent is mostly the same, though there is some ambiguity about the role of functionality in infringement.³⁸ Copyright has relatively few general defenses because its threshold for validity is so low, though as we discuss below, some courts regard certain aspects of validity (like merger) as defenses. On the other hand, copyright is chock full of specific exceptions and defenses that protect particular users in particular situations.³⁹ And even cases that seem from some vantage point to be concerned with validity look that way only because of the nature of the defendant's use—which is to say that they might really be defendant-specific scope decisions after all.

Take the famous case of *Baker v. Selden*, which is widely regarded as the source of copyright's idea-expression doctrine.⁴⁰ According to the Court, Selden's attempt to enforce rights against Baker for the latter's use of accounting forms that were substantially similar to those illustrated in Selden's book amounted to an illegitimate attempt to use copyright to prevent use of an unpatented method of accounting.⁴¹ As the Court noted, "the mere copyright of Selden's book did not confer upon him the exclusive right to make and use account-books, ruled and arranged as designated by him and described and illustrated in said book."⁴²

Because the Court's opinion speaks in terms of protectability, it is natural to see the case in terms of validity.⁴³ But Selden's forms

37. See *Madey v. Duke Univ.*, 307 F.3d 1351, 1360-61 (Fed. Cir. 2002) (holding the experimental use defense inapplicable in virtually all circumstances). More recently, the America Invents Act expanded a prior user right defense somewhat, 35 U.S.C. § 273, but that defense has so far not been applied in a reported decision.

38. See *Richardson v. Stanley Works, Inc.*, 597 F.3d 1288, 1293-94 (Fed. Cir. 2010).

39. See, e.g., 17 U.S.C. §§ 107-122.

40. See 101 U.S. 99, 103 (1879).

41. See *id.* at 105 ("The description of the art in a book, though entitled to the benefit of copyright, lays no foundation for an exclusive claim to the art itself. The object of the one is explanation; the object of the other is use. The former may be secured by copyright. The latter can only be secured, if it can be secured at all, by letters-patent.")

42. *Id.* at 107.

43. See *id.* ("The conclusion to which we have come is, that blank account-books are not the subject of copyright.")

were included in a book that clearly was the subject of copyright, and the question for the Court was whether Baker's creation of similar account books infringed the copyright in Selden's book.⁴⁴ To put it in our terms, the issue was the scope of Selden's rights in his book. And the Court determined that those rights didn't allow Selden to prevent use of the forms for the purpose of practicing the art⁴⁵—itself a controversial conclusion at the time.⁴⁶ That wasn't the end of the story, however, as the Court strongly implied that other uses of the forms might have fallen within Selden's rights. As the Court said:

[W]here the art [the book] teaches cannot be used without employing the methods and diagrams used to illustrate the book, or such as are similar to them, such methods and diagrams are to be considered as necessary incidents to the art, and given therewith to the public; *not given for the purpose of publication in other works explanatory of the art, but for the purpose of practical application.*⁴⁷

Baker v. Selden, then, isn't really a validity case or an infringement case. It is about the proper scope of a valid copyright.

A number of other famous copyright cases also are really about the scope of rights. In *Lotus Development Corp. v. Borland International, Inc.*, for instance, the district court held that Borland infringed Lotus's copyright in its spreadsheet program by copying the menu commands that activated features in the Lotus program so that Borland users could port their spreadsheets over from Lotus and use familiar command keys.⁴⁸ For the district court, Lotus had a valid copyright in its program (surely true), so the only question

44. *See id.* at 99-100.

45. *See id.* at 103 ("The copyright of a book on perspective, no matter how many drawings and illustrations it may contain, gives no exclusive right to the modes of drawing described, though they may never have been known or used before. By publishing the book, without getting a patent for the art, the latter is given to the public.")

46. For a general discussion of the history and context of the case, see Pamela Samuelson, *The Story of Baker v. Selden: Sharpening the Distinction Between Authorship and Invention*, in *INTELLECTUAL PROPERTY STORIES* 159 (Jane C. Ginsburg & Rochelle Cooper Dreyfuss eds., 2006).

47. *Baker*, 101 U.S. at 103 (emphasis added).

48. *See* 49 F.3d 807, 812 (1st Cir. 1995).

was one of infringement.⁴⁹ The court of appeals reversed. In the majority's view, the key question was validity, but it determined that the lawful scope of Lotus's copyright did not extend to menu commands because they were unprotectable methods of operation.⁵⁰ The concurrence, by contrast, would have treated the menu commands as part of the valid copyright but would have treated Borland's use of those commands for compatibility purposes as a privileged use.⁵¹ The difference in approach is a function of scope, and the consequences differ significantly. A validity approach bars all protection for the work in question (though note that even here the *Lotus* majority engaged in a scope inquiry by focusing only on the protectability of a particular part of the program, while the district court focused on the program as a whole), whereas a fair use defense privileges some but not all uses of that work, leaving others as infringing. Other copyright cases, particularly those involving software, produce similar debates about the proper approach.⁵²

Trademark is the most hybrid of the regimes. Some defenses—including genericness, abandonment, and functionality (at least mechanical functionality)—are really validity doctrines with all-or-nothing effect. Other doctrines, like descriptive fair use and the first sale doctrine, are defendant-specific.⁵³ Still other doctrines—*Rogers*

49. *See id.* at 811.

50. *See id.* at 816. For a defense of this ruling pointing out that 17 U.S.C. § 102(b) excludes much more than ideas from copyright protection, see Pamela Samuelson, *Why Copyright Law Excludes Systems and Processes from the Scope of Its Protection*, 85 TEX. L. REV. 1921, 1921-23 (2007).

51. *See Lotus*, 49 F.3d at 821-22 (Boudin, J., concurring).

52. *See, e.g., Oracle Am., Inc. v. Google, Inc.*, 750 F.3d 1339, 1347-48 (Fed. Cir. 2014); *Am. Dental Ass'n v. Delta Dental Plans Ass'n*, 126 F.3d 977, 979 (7th Cir. 1997). In *Oracle*, the district court held that Oracle's Java applications program interfaces (APIs)—bits of code designed to allow other programs to interact with Java—were unprotectable methods of operation that Google was free to copy. *See Oracle*, 750 F.3d at 1347-48. The Federal Circuit reversed, holding that the APIs were independently protectable and that copying them was infringement, but leaving open the question of whether doing so was a fair use. *See id.* at 1348.

53. As one of us has previously emphasized, trademark law is plagued with uncertainty about the extent to which various defensive doctrines are really defenses in the sense that they are distinct from the prima facie case of infringement. *See* William McGeeveran & Mark P. McKenna, *Confusion Isn't Everything*, 89 NOTRE DAME L. REV. 253, 277 (2013). That uncertainty has significant consequences for scope, in that it is often unclear whether the "defenses" are doing any work to limit scope or instead are simply describing situations under which confusion is unlikely and therefore noninfringing. *See* Gideon Parchomovsky & Alex Stein, *Intellectual Property Defenses*, 113 COLUM. L. REV. 1483, 1503-04 (2013) (asserting that

v. Grimaldi and related expressive use doctrines, for example—are defendant-specific but perhaps not really defenses as much as expressions of what falls outside the scope of a trademark owner’s prima facie legal rights.⁵⁴

B. Why Scope Matters Now

Doctrines of infringement, validity, and defenses have been around almost as long as IP law.⁵⁵ Why, then, has the issue of scope become important now? Three developments in IP law over the course of the twentieth century explain why scope is a much more important issue today than it was in the past. First, most areas of IP have loosened their infringement standards. In copyright, for example, not only did the law recognize new exclusive rights, but courts watered down the level of similarity necessary for a defendant’s work to be deemed infringing. Whereas once a defendant’s work had to be highly similar to the plaintiff’s work as a whole, courts now find infringement based on the similarity of the defendant’s work to small parts of the plaintiff’s work,⁵⁶ and even the similarity of small parts of the defendant’s work to small parts of the plaintiff’s work.⁵⁷ Overall similarity is no longer necessary; fragmented similarity is enough.⁵⁸

“defenses play an important role in delineating the scope of protection that our law affords to trademarks”). For a full discussion of copyright “limitations and exceptions” that spans both true defenses and limitations that are nonetheless part of the prima facie case of copyright infringement, see Pamela Samuelson, *Justifications for Copyright Limitations & Exceptions*, in *COPYRIGHT LAW IN AN AGE OF LIMITATIONS AND EXCEPTIONS* (Ruth Okediji ed., forthcoming 2016) (manuscript at 1-4), <http://ssrn.com/abstract=2476669> [<https://perma.cc/NX82-T9PM>].

54. Under *Rogers v. Grimaldi*, which now seems to be the dominant approach to expressive-use cases, use of a mark in or in relation to an expressive work is not infringing so long as it has some artistic relevance to the work and does not explicitly mislead. 875 F.2d 994, 1005 (2d Cir. 1989). This is clearly a scope doctrine, but—to our point about the superficiality of these categories—probably more a question of infringement than defense.

55. Though, as we note below, modern courts conceive of the distinctions between these phases much more rigidly than their predecessors.

56. See *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 569 (1985).

57. The music sampling cases are the most extreme examples here. See *Bridgeport Music, Inc. v. Dimension Films*, 410 F.3d 792, 799-805 (6th Cir. 2005). But the phenomenon is not limited to sound recordings. See *Oracle*, 750 F.3d at 1347-48; *ABKCO Music, Inc. v. Harrisongs Music, Ltd.*, 722 F.2d 988, 997 (2d Cir. 1983).

58. See Justin Hughes, *Size Matters (or Should) in Copyright Law*, 74 *FORDHAM L. REV.* 575, 618 (2005).

It's possible to see this evolution in terms of how to define the work with which the defendant's work should be compared rather than as a watering down of the similarity standard. Instead of defining works in terms of the tangible artifacts in which expression was fixed (for instance, a book, a song, or a motion picture), copyright law now conceives of the unit of protection in terms of originality. As a result, rather than conceiving of the plaintiff's work as an entire book or song, as they once would have, courts now are willing to look inside the plaintiff's work for small increments of original expression. As Justin Hughes persuasively described, this was the predictable result of copyright's de-emphasis of registration and abstraction away from tangible objects to intangible "works" as the unit of protection.⁵⁹ Until registration was eliminated as a requirement for copyright protection, "the system largely policed itself in terms of the size and concept of 'works' which creators and owners deemed worthy of protection."⁶⁰ But as the uproar over the Ninth Circuit's decision in *Garcia v. Google, Inc.* demonstrates, when only originality delineates a copyrighted work, courts have little meaningful guidance on how to draw the boundaries of an "original work of authorship."⁶¹

Whether we see the development of copyright's infringement standard as a lessening of the similarity standard or an increasing willingness to define works granularly, the point is that much smaller bits of similarity are enough to trigger liability. And that puts much more pressure on the infringement apparatus to ensure that the similarity relates to protectable matter.

Trademark law experienced something similar. In the early part of the twentieth century, trademark infringement and unfair competition, while doctrinally distinct, were both concerned with the same wrong of trade diversion.⁶² As a result, only directly competing uses were actionable, as those are the only uses that could result in passing off (selling the defendant's goods as if they were the plaintiff's

59. *Id.* at 580; see also Paul Goldstein, *What Is a Copyrighted Work? Why Does It Matter?*, 58 UCLA L. REV. 1175, 1178 (2011); Madison, *supra* note 20, at 2-3.

60. Hughes, *supra* note 58, at 580.

61. See *Garcia v. Google, Inc.*, 766 F.3d 929 (9th Cir. 2014) (finding that actress had a "copyright interest" in her performance within a film), *rev'd en banc*, 786 F.3d 733, 749-50 (9th Cir. 2015) (rejecting plaintiff's claim that her performance was copyrightable work).

62. Mark P. McKenna, *The Normative Foundations of Trademark Law*, 82 NOTRE DAME L. REV. 1839, 1849 (2007).

goods). But courts dramatically expanded the range of actionable confusion over the course of the twentieth century. No longer was it necessary for consumers to believe that the defendant's goods were those of the plaintiff, or even that the plaintiff actually produced those goods. Instead, it was enough that consumers believed there was some relationship between the plaintiff and defendant—that the plaintiff “sponsored” the defendant or was somehow “affiliated” with it.⁶³

Whatever the merits of this expansion, it naturally reduced the level of similarity necessary to find infringement. When courts sought only to prevent trade diversion, they were concerned only with uses of a mark for directly competing goods. Infringement analysis therefore consisted simply of a comparison of the parties' respective marks; no likelihood of confusion factors were necessary.⁶⁴ And since the comparison of the marks was meant only to determine whether consumers were likely to believe that the defendant's goods *were* the plaintiff's goods, only highly similar uses could be found infringing.⁶⁵ Once courts could find infringement by concluding only that consumers might believe there was some connection between the parties, less similarity between the marks themselves was necessary.⁶⁶ For the same reason, the goods marked no longer needed to be the same.⁶⁷

63. We have previously written on the problems with overbroad notions of “sponsorship” or “affiliation.” See Mark A. Lemley & Mark McKenna, *Irrelevant Confusion*, 62 STAN. L. REV. 413, 427 (2010).

64. See McKenna, *supra* note 62, at 1904 (noting that courts only developed the likelihood of confusion factors after jettisoning the requirement of direct competition). For a more thorough account of the development of the likelihood of confusion factors, see Robert G. Bone, *Taking the Confusion out of “Likelihood of Confusion”*: Toward a More Sensible Approach to Trademark Infringement, 106 NW. U. L. REV. 1307, 1316-36 (2012).

65. Unfair competition cases obviously didn't involve comparison of trademarks, since by definition they didn't involve technical trademarks. But the same general point held in those cases. Unfair competition required a finding that, despite not using a trademark, the defendant's acts were likely to deceive consumers into thinking that the defendant's goods were those of the plaintiff. And that finding required use of something highly similar to what the plaintiff was using—whether use of a similar trade name, geographic term, or product packaging—in a way that was likely to deceive.

66. Cf. Rebecca Tushnet, *Looking at the Lanham Act: Images in Trademark and Advertising Law*, 48 HOUS. L. REV. 861, 888-89 (“If the question were only ‘when you look at X, do you think it is Y?’ then it would be much easier to avoid defining X. With the question ‘when you look at X, do you associate it with Y?’ there is a greater premium on knowing exactly what X is—how far something can deviate from X before it no longer reminds consumers of X.”).

67. For a discussion, see Mark A. Lemley & Mark P. McKenna, *Owning Mark(et)s*, 109

Loosening of the infringement standard in design patent is more recent, and possibly less consistent, but we think equally notable. The basic rule of infringement has been the same since *Gorham Co. v. White*, under which a design patent is infringed “if, in the eye of an ordinary observer, giving such attention as a purchaser usually gives, two designs are substantially the same, if the resemblance is such as to deceive such an observer, inducing him to purchase one supposing it to be the other.”⁶⁸ While this standard has never required absolute identity, as the finding of infringement in *Gorham* itself indicates, for most of their history, design patents were regarded as extremely narrow. Indeed, Judge Rich once famously wrote that “[d]esign patents have almost no scope.”⁶⁹ To whatever extent Judge Rich was right at the time he wrote, it is much harder to describe design patents the same way today.

Second, this relaxation of the infringement standards happened at the same time courts and Congress were significantly expanding the subject matter of nearly every area of IP. We mean this both in the sense that IP regimes included new types of protectable subject matter (in copyright, software, sound recordings, architectural works, to name a few), and also in the sense that new technologies operated on existing subject matter categories to effectively create new subject matter (for example, software patents, or design patents on graphical user interfaces). In trademark law this subject matter expansion has been particularly dramatic. Whereas only technical trademarks were previously eligible for registration and enforcement under federal law, modern law recognizes “anything ... capable of carrying [source] meaning” as a potential trademark.⁷⁰

MICH. L. REV. 137, 146-56 (2010).

68. 81 U.S. 511, 528 (1871).

69. *In re Mann*, 861 F.2d 1581, 1582 (Fed. Cir. 1988). In fact, precisely because they perceived design patents to have no scope, many commentators who were interested in broader protection for design focused on expanded copyright protection as a solution. *See, e.g.*, Ralph S. Brown, *Design Protection: An Overview*, 34 UCLA L. REV. 1341, 1356 (1987) (calling design patent “a Cinderella who never goes to the ball”); J.H. Reichman, *Design Protection After the Copyright Act of 1976: A Comparative View of the Emerging Interim Models*, 31 J. COPYRIGHT SOC’Y 267, 350 & n.475 (1983) (calling the failure to achieve comprehensive short-term protection for the design of useful articles a “mutilation” that left copyright with the separability test and no effective protection for the designs excluded from copyright as a result).

70. *Qualitex Co. v. Jacobson Prods. Co.*, 514 U.S. 159, 162 (1995).

This has collapsed unfair competition entirely into trademark law, bringing into trademark law subject matter that once would have been subject only to unfair competition enforcement. As a result, trademark infringement questions now often involve much more complex comparisons. Rather than simple comparisons of word marks or logos, cases now regularly involve comparisons of advertising slogans, product packaging, and even the product designs themselves. And though even comparisons of word marks present scope issues, the difficulty of those questions is compounded with more complex marks because those marks are likely to include much more unprotectable matter.

Combining these developments, we can see why the kind of scope questions with which we are concerned are much more pressing now. If courts require a very high level of similarity to find infringement, they can afford to be less worried about the extent to which the similarity between the works is due to unprotectable elements. To take a highly stylized example, if, in order to find infringement, a court must find the defendant's work, trademark, or design 90 percent similar to the plaintiff's, then it can be less concerned about the fact that 75 percent of the plaintiff's work is unprotectable. In order to reach the 90 percent similarity threshold, the defendant will necessarily have copied protectable matter. If, on the other hand, the defendant's work must only be 65 percent similar to be deemed infringing, then it is a serious concern if 75 percent of the plaintiff's work is unprotectable, because the 65 percent threshold could easily be met entirely by similarity to unprotectable matter.

Most areas of IP put pressure on both sides of this equation over the twentieth century—lowering the level of similarity necessary to find infringement *and* protecting works with higher proportions of unprotectable elements. To manage the resulting problems, courts in each area of IP have developed elaborate doctrinal structures for evaluating validity and infringement. These doctrines are supposed to limit the scope of a right to correspond to the point of novelty, originality, or other elements of protectibility. And they stand in some contrast to their historical predecessors. Patent, design patent, and trademark law all paid much less attention to the validity-infringement distinction in the nineteenth century than they do today. Cases tended to be resolved by concluding that the defendant was or was not liable, rather than separating the decision into

validity and infringement grounds.⁷¹ In part, that was because invalidation of a patent did not have the same consequence it has today: eliminating the plaintiff's right to sue anyone else for infringement.⁷² We only needed a way to assess scope once we separated validity from infringement clearly and began applying different standards to each.

But the development of these safeguards has run into conflict with the third development in the twentieth century—the increasing role of juries in resolving infringement and validity disputes. Jury trials expanded dramatically in patent cases, from less than 2.5 percent of all trials in the middle of the twentieth century to nearly 70 percent by its end.⁷³ Courts in the mid-twentieth century expanded the role of the jury and its ordinary observer standard in assessing copyright infringement.⁷⁴ Complex rules for deciding the proper scope of an IP right are difficult things to convey to the jury.

II. THE PROBLEMS WITH SEPARATION

A. *Gaps in the Fabric of IP Scope*

Notwithstanding the interconnectedness of the various scope questions courts must address, those questions arise in different parts of the tripartite case structure typically associated with IP cases. The result of this separation between validity, infringement, and defenses is a number of disconnects—gaps in the IP regime through which either an IP owner or an infringer can sneak. By gaps, we mean circumstances in which the law should reach a consistent scope result but fails to do so because each IP doctrine assumes that others will take care of the problem.

71. See *Smith v. Whitman Saddle Co.*, 148 U.S. 674, 682 (1892).

72. See Mark A. Lemley, *Why Do Juries Decide If Patents Are Valid?*, 99 VA. L. REV. 1673, 1682 (2013) (discussing this change).

73. For a discussion of this history, see *id.* at 1704-06.

74. See *Arnstein v. Porter*, 154 F.2d 464, 468 (2d Cir. 1946).

B. Gaps Exist in the Fabric of IP for a Number of Reasons

1. Different Decisionmakers

Sometimes different actors decide different doctrines, often at different times. Some IP issues are decided by judges, others by juries.⁷⁵ Judges and juries can have fundamentally different approaches to similar questions.⁷⁶ When a judge resolves one issue (say, granting summary judgment that a patent is infringed) and declines to resolve others (whether the patent is invalid) on the assumption that the jury will handle the latter issue, the fact that the jury rather than the judge makes the second determination leaves open the possibility that the two issues will not be resolved consistently. That is even more likely when the jury hears only part of the story because the judge has already decided other parts. In our example, a jury that might be inclined to decide that a plaintiff's patent shouldn't cover the defendant's conduct doesn't have an opportunity to do so because the judge has already resolved the infringement issue. The jury will hear only the all-or-nothing question of whether the patent is invalid without learning of the dispute over how broadly it is being construed. As we noted in the last section, this is a bigger problem now than it was in the past, not only because the division between infringement and validity is sharper, but also because juries play more of a role in assessing validity than they used to.

For many IP regimes, the problem of different decisionmakers is heightened by the fact that the U.S. Patent and Trademark Office (PTO) is set up to resolve some IP issues—specifically, validity questions—but not others. Thus, the PTO decides the validity of the patents and trademarks it considers, initially or in reexamination,

75. For a discussion on the division between judge and jury in patent cases, along with some skepticism as to whether juries are required to decide many of these issues, see Eileen M. Herlihy, *The Ripple Effect of Seventh Amendment Decisions on the Development of Substantive Patent Law*, 27 SANTA CLARA COMPUTER & HIGH TECH. L.J. 333, 343 (2011). See generally Lemley, *supra* note 72. For consideration of a thorny issue on the allocation of decisionmaking in trademark cases, see generally *Hana Fin., Inc. v. Hana Bank*, 135 S. Ct. 907 (2015) (holding that continuing commercial impression, the standard for tacking, is a fact question which generally should be decided by a jury).

76. See generally Fromer & Lemley, *supra* note 25 (describing the differences between an expert and a consumer analysis).

without having to confront the scope of the resulting right, or considering who might be sued for infringing it and whether she might have a defense to suit.⁷⁷ Indeed, the PTO even applies different substantive standards than the courts on important issues like claim construction in patent law⁷⁸ and likelihood of confusion in trademark law.⁷⁹

In patent law, that structural disconnect has recently become more pronounced. Whereas the PTO considers only validity issues when deciding to issue a patent, courts traditionally have considered both validity and infringement issues when the patent is enforced. Beginning in 1980, however, the PTO got limited power to reexamine its earlier decision to issue a patent.⁸⁰ Those powers were expanded dramatically in 2011 by the America Invents Act, which gave the PTO the power to conduct inter partes review,⁸¹ postgrant oppositions,⁸² and covered business method reviews.⁸³ Patent challengers have instituted thousands of such reviews since 2011.⁸⁴ And for one of these proceedings, covered business method review, courts are generally expected to stay litigation proceedings until the PTO review process is complete.⁸⁵ The result is that a growing number of

77. See generally Sarah Tran, *Patent Powers*, 25 HARV. J.L. & TECH. 609 (2012) (discussing the powers of the PTO).

78. While courts seek to determine the proper meaning of a term to a person of ordinary skill in the art, the PTO applies the “broadest reasonable construction” of a term, not only when granting patents in the first instance but even when reexamining issued patents that are also pending in litigation. See generally Dawn-Marie Bey & Christopher A. Cotropia, *The Unreasonableness of the Patent Office’s “Broadest Reasonable Interpretation” Standard*, 37 AIPLA Q.J. 285 (2009) (discussing and criticizing this difference).

79. The PTO evaluates likelihood of confusion considering only the similarity of the marks as depicted, not taking into account the context in which the mark is used. See 3 MCCARTHY, *supra* note 23, § 20:15 (“In determining likelihood of confusion in an opposition, it is the mark as shown in the application and as used on the goods described in the application which must be considered, not the mark as actually used by applicant.”). The difference between PTO and judicial likelihood of confusion standards is the subject of a recent Supreme Court case. See *B&B Hardware, Inc. v. Hargis Indus., Inc.*, 135 S. Ct. 1293 (2015).

80. See 35 U.S.C. §§ 301-304 (2012).

81. See *id.* § 311.

82. See *id.* § 321.

83. See *id.* note (b)(2) (Transitional Program for Covered Business Method Patents).

84. Office of Patent Legal Admin., *AIA Statistics*, U.S. PAT. & TRADEMARK OFF. (Mar. 27, 2015), <http://www.uspto.gov/patent/laws-and-regulations/america-invents-act-aia/aia-statistics> [<https://perma.cc/SCL8-ZTWE>].

85. See 35 U.S.C. § 321 note (b)(2) (Transitional Program for Covered Business Method Patents); *VirtualAgility Inc. v. Salesforce.com*, 759 F.3d 1307, 1310 (Fed. Cir. 2014).

patent cases will be resolved by considering the question of validity in the abstract, without infringement and defenses being presented to the PTO.⁸⁶ While there is some risk the PTO will wrongly invalidate a patent claim that should properly have been narrowed, there is also a risk that it will hold a patent claim valid without considering the breadth of that claim, and that such a decision will influence district courts in subsequent litigation where the breadth question is actually presented.⁸⁷

2. *Different Burdens of Proof*

Even if the decisionmaker is the same, validity, infringement, and defenses often come with different burdens of proof. For example, registered IP rights are all presumed valid when challenged in court; the standard of proof differs from a preponderance of the evidence in copyright and trademark cases to clear and convincing evidence in patent and design patent cases. Proving infringement is the IP owner's burden, always by preponderance of the evidence; proving a defense is a defendant's burden, and again the level of proof required may differ by doctrine and defense.⁸⁸

86. On the uneven growth of the administrative powers of the PTO over patent law, see, for example, Sapna Kumar, *Expert Court, Expert Agency*, 44 U.C. DAVIS L. REV. 1547 (2011); Jacob S. Sherkow, *Administering Patent Litigation*, 90 WASH. L. REV. 205, 218-19 (2015); Sarah Tran, *Administrative Law, Patents, and Distorted Rules*, 80 GEO. WASH. L. REV. 831, 855-58 (2012); Tran, *supra* note 77, at 609, 626-30.

The problem is even worse in Europe, where the infringement and validity decisions are fully separated, with a court deciding infringement and the patent agencies deciding validity. That has led to a number of cases in which courts grant injunctions on the basis of patents that are later determined to be invalid. See generally Katrin Cremers et al., *Invalid but Infringed? An Analysis of Germany's Bifurcated Patent Litigation System* (Max Planck Inst. for Innovation and Competition Research Paper Series, Paper No. 14-14, 2015), <http://ssrn.com/abstract=2504507> [<https://perma.cc/JV55-FWNN>] (discussing instances of patents found invalid after a determination of infringement).

87. Courts seem increasingly willing to defer to decisions of administrative courts in the PTO. See, e.g., *B & B Hardware, Inc. v. Hargis Indus., Inc.*, 135 S. Ct. 1293, 1299 (2015) (holding that rulings of the Trademark Trial and Appeal Board on likelihood of confusion collaterally estopped litigating parties in courts).

88. Inequitable conduct in patent law, for instance, requires proof by the defendant by clear and convincing evidence. See *Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1287 (Fed. Cir. 2011) (en banc). Most other defenses require a defendant to show only a preponderance of the evidence. See Roger Allan Ford, *Patent Invalidity Versus Noninfringement*, 99 CORNELL L. REV. 71, 103 (2013).

Even the same doctrine may require different standards of proof for different parties and at different times. Patent invalidity must be proven by clear and convincing evidence in court, but only by a preponderance of the evidence during patent examination or during some—but not all—forms of reexamination in the PTO.⁸⁹ The fact that someone is the first inventor of a patent must be proved in an interference proceeding by a preponderance of the evidence if the patent has not yet issued or been published, but by clear and convincing evidence if the challenge is raised thereafter.⁹⁰ The Supreme Court cemented this distinction in *Commil USA, LLC v. Cisco Systems, Inc.*, declaring that “[v]alidity and infringement are distinct issues, bearing different burdens, different presumptions, and different evidence.”⁹¹

The fact that different parts of the IP fabric are subject to different standards of proof leaves open the possibility that those burdens will distort the overall result in the case. If there is a preponderance of the evidence supporting a patent defendant, but not clear and convincing evidence, whether we treat the evidence as relevant to validity or to infringement will determine who wins the case. Similarly, when evidence is hard to come by, classifying a doctrine as part of infringement (which the plaintiff must prove) or a defense (which the defendant must prove) will determine the outcome. Further, as Jonathan Masur and Lisa Ouellette have shown, the difference in deference regimes can lead to systematic mistakes in applying the law from one context to another.⁹²

89. See, e.g., *In re Swanson*, 540 F.3d 1368, 1377-78 (Fed. Cir. 2008).

90. See *Bruning v. Hirose*, 161 F.3d 681, 684-86 (Fed. Cir. 1998). No challenge at all can be made more than a year after the patent is issued or the application is published. 35 U.S.C. § 135(a)(2) (2012) (abrogated prospectively by the America Invents Act of 2011).

91. 135 S.Ct. 1920, 1929 (2015) (quoting *Commil USA, LLC v. Cisco Sys., Inc.*, 720 F.3d 1361, 1374 (Fed. Cir. 2013)). For an argument that this separation in patent law owes its origin to the pre-Federal Circuit separation of validity and infringement determinations into different courts, see Jeffrey A. Lefstin, *The Constitution of Patent Law: The Court of Customs and Patent Appeals and the Shape of the Federal Circuit's Jurisprudence*, 43 LOY. L.A. L. REV. 843, 874-83 (2010).

92. See generally Jonathan S. Masur & Lisa Larrimore Ouellette, *Deference Mistakes*, 82 U. CHI. L. REV. 643 (2015) (discussing consequences of deference mistakes).

3. *A Reluctance to Invalidate IP Rights*

Courts, and particularly juries, are reluctant to second-guess the PTO and declare IP rights invalid altogether. Roger Ford has argued that courts are accordingly too willing to resolve cases on infringement grounds even when the IP right should, in truth, be invalid.⁹³ In fact, however, this reluctance often makes sense, not so much because the PTO has the final say, but because many times an IP right is properly valid but narrow. But a low threshold for validity only makes sense if we have confidence that we can sort out efforts to overclaim the scope of that right using infringement doctrines. And often we don't have the proper tools in infringement or defenses. When validity and infringement are separated, the reluctance of courts to invalidate an IP right altogether means that they find it hard to reach the right result when the tools for limiting the right are classed as invalidity or general defenses rather than infringement doctrines or conduct-specific defenses. This is true of the functionality doctrine in both trademark and design patents, for instance. Because we don't have a clear defense for nontrademark uses⁹⁴—uses of a mark by a defendant, not for purposes of signaling source, but to take advantage of the intrinsic characteristics of the product—courts are often faced with a choice between invalidating a mark altogether or letting the plaintiff get away with expanding the scope of her right beyond its proper bounds.⁹⁵ Courts in this situation tend to expand the right rather than invalidate it,⁹⁶ though

93. See generally Ford, *supra* note 88.

94. *Rescuecom Corp. v. Google Inc.*, 562 F.3d 123 (2d Cir. 2009), is generally considered responsible for effectively killing the trademark use doctrine, though it remains good law in several circuits. For discussion of the state of the doctrine and its merits, compare Stacey L. Dogan & Mark A. Lemley, *Grounding Trademark Law Through Trademark Use*, 92 IOWA L. REV. 1669, 1670 (2007) (defending the importance of the doctrine), with Graeme B. Dinwoodie & Mark D. Janis, *Confusion Over Use: Contextualism in Trademark Law*, 92 IOWA L. REV. 1597, 1599-603 (2007) (arguing that there should be no limitation on the sort of conduct that could infringe trademarks, if found to confuse consumers), and Mark P. McKenna, *Trademark Use and the Problem of Source*, 2009 U. ILL. L. REV. 773, 774-76 (challenging the ability of courts to define what constitutes a trademark use).

95. Cf. Chao, *supra* note 8, at 1359 ("Although there are numerous existing doctrines that try to prevent claims from straying too far from their specification, these doctrines offer binary outcomes ill suited for patent law.")

96. See *Rosetta Stone Ltd. v. Google, Inc.*, 676 F.3d 144, 162 (4th Cir. 2012) (holding that the functionality doctrine applies only to the mark as used by the plaintiff and not to the defendant's use of the mark); *Au-Tomotive Gold, Inc. v. Volkswagen of Am., Inc.*, 457 F.3d

sometimes they have made up new defenses or other doctrines to try to avoid the choice.⁹⁷

4. *Strategic Behavior*

Each of the problems we just identified exists even if no one tries to exploit these gaps. They are inherent disconnects between the rules of validity, infringement, and defenses. But in fact parties inevitably seek to exploit them for their own advantage. Parties can exploit the gaps in a variety of ways. One of the simplest is trying to separate the determination of validity from the determination of infringement, either by having them done at separate times or by different decisionmakers. Whenever the two are separated, IP owners will argue in the infringement proceedings that their right is quite broad, covering what the defendant does, only to turn around and argue in validity proceedings that their right is quite narrow and therefore valid. Accused infringers will do the reverse. Without a single integrated scope proceeding, there is no easy way to constrain this behavior.

C. *IP Doctrines and the Nose of Wax*

The result of these gaps has been that every IP regime has instances in which courts systematically reach the wrong decisions on the scope of the right. In this Section, we offer a few examples from each regime.

1062, 1072-74 (9th Cir. 2006) (refusing to apply the aesthetic functionality doctrine to the defendant's use of the marks for fear of invalidating those marks).

97. In *Christian Louboutin S.A. v. Yves Saint Laurent American Holding, Inc.*, 696 F.3d 206 (2d Cir. 2012), for instance, the court evaded the aesthetic functionality doctrine and limited Louboutin's rights to contrasting red-sole shoes so as to find for the defendant but avoid invalidating the mark altogether or even explicitly deciding whether consumers were confused. *See id.* at 212. And several courts have sought to apply a doctrine of "defensive functionality" to protect the defendant's use from liability without invalidating the mark altogether, though those efforts have ultimately been rejected. *See, e.g.*, *Rosetta Stone Ltd. v. Google, Inc.*, 730 F. Supp. 2d 531 (E.D. Va. 2010), *rev'd in relevant part*, 676 F.3d 144, 161 (4th Cir. 2012); *Fleischer Studios, Inc. v. A.V.E.L.A., Inc.*, 636 F.3d 1115, 1124 (9th Cir. 2011), *rev'd on reh'g*, 654 F.3d 958 (9th Cir. 2011). One of us has previously argued that those courts were on the right track. Mark P. McKenna, (*Dys*)*Functionality*, 48 HOUS. L. REV. 823, 854-57 (2011).

1. Copyright

Copyright law is the least reliant of the IP regimes on validity doctrines and therefore the most reliant on infringement doctrines and defenses to manage the scope of a party's rights. Copyright's originality threshold requires only independent creation and a minimal level of creativity.⁹⁸ As the Supreme Court said in *Feist Publications v. Rural Telephone Service*, "the requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, no matter how crude, humble or obvious it might be."⁹⁹ Copyright thus allows protection for works that are extremely similar to works in the public domain or other copyrightable works,¹⁰⁰ or that incorporate many unprotectable elements.

Copyright tries to deal with this fact through a number of doctrines that limit the scope of protection. Copyright does not extend to facts or ideas,¹⁰¹ and while copyright can subsist in the selection and arrangement of facts or ideas, courts will deny protection in cases in which idea and expression merge because there are few ways to express an idea.¹⁰² Copyright also does not extend to "scènes-à-faire"—stock scenes that naturally flow from a common theme.¹⁰³ Section 102(b) of the Copyright Act is in some senses a

98. See *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 345 (1991) ("Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity.").

99. *Id.* (quoting 1 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 1.08 [c][1] (1990)).

100. Indeed, in principle a work is protectable even if it is *identical* to prior works, so long as the work was independently created. See *Sheldon v. Metro-Goldwyn Pictures Corp.*, 81 F.2d 49, 54 (2d Cir. 1936) ("Borrowed the work must indeed not be, for a plagiarist is not himself pro tanto an 'author'; but if by some magic a man who had never known it were to compose anew Keats's Ode on a Grecian Urn, he would be an 'author,' and, if he copyrighted it, others might not copy that poem, though they might of course copy Keats's.").

101. *Feist*, 499 U.S. at 344-45.

102. See, e.g., *Morrissey v. Procter & Gamble Co.*, 379 F.2d 675, 678-79 (1st Cir. 1967) (finding *Morrissey's* rules for a promotional contest uncopyrightable on merger grounds).

103. See, e.g., *Walker v. Time Life Films, Inc.*, 784 F.2d 44, 50 (2d Cir. 1986) (finding no protection for common elements in police fiction, such as "drunks, prostitutes, vermin and derelict cars" and "[f]oot chases and the morale problems of policemen, not to mention the familiar figure of the Irish cop"); *Evans v. Wallace Berrie & Co.*, 681 F. Supp. 813, 817 (S.D. Fla. 1988) ("Such similarities as using a sand dollar as currency, foods made of seaweed, seahorses for transportation and plates made of oyster or mother of pearl are not protected

codification of the idea/expression dichotomy, but that section goes further and excludes protection for processes, systems, and methods of operation.¹⁰⁴

These are important doctrines that sometimes operate at the validity stage to preclude protection under any circumstances.¹⁰⁵ In *Feist*, for example, the Court found Rural Telephone's white pages uncopyrightable because the names, towns, and telephone numbers were unprotectable facts, and Rural had not selected or arranged those facts in an original way.¹⁰⁶ As a result, Feist's white pages could not infringe, even though Feist had actually copied a number of listings from Rural's phone book.¹⁰⁷ And in *Morrissey v. Procter & Gamble*, the court refused protection to Morrissey's sweepstakes rules on merger grounds even though it found that "there was more than one way of expressing" the rules for the game, and that there was "almost precise similarity" between the plaintiff's and the defendant's rules.¹⁰⁸

But, in general, copyright law only rarely disqualifies works altogether and instead generally relies on infringement doctrine to limit the scope of rights in a work so that it reflects that work's marginal copyrightability. Indeed, even the doctrines we just identified often don't fit neatly in the validity box. Courts most commonly recite these restrictions not to deny the plaintiff protection at all,

similarities of expression, but are more accurately characterizations that naturally follow from the common theme of an underwater civilization.").

104. 17 U.S.C. § 102(b) (2012); see Samuelson, *supra* note 50, at 1944-61 (demonstrating that § 102(b) was intended to codify the exclusion of processes, systems, and methods of operation, in keeping with case law developed following *Baker v. Selden*).

105. See Pamela Samuelson, *Reconceptualizing Copyright's Merger Doctrine* 16-19 (unpublished manuscript) (on file with authors) (demonstrating that merger sometimes functions as a validity doctrine or a limitation on scope and not just as a defense to infringement); see also 1 PAUL GOLDSTEIN, GOLDSTEIN ON COPYRIGHT § 2.3.2, at 2:38.1 (3d ed. Supp. 2013-1) ("As traditionally conceived, the merger doctrine determines whether the work is copyrightable; if the idea and expression of a jeweled bee pin merge, the pin's creator has no copyright."). The U.S. Copyright Office's Compendium of Copyright Office Practices confirms this view, identifying merger as one of the bases on which the Office may refuse registration. UNITED STATES COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 313.3(B) (3d ed. 2014).

106. See 499 U.S. at 361.

107. See *id.* at 343-44 (noting that Feist copied 1309 listings that were identical to listings in Rural's phone books, including four fictitious listings).

108. 379 F.2d 675, 678 (1st Cir. 1967). As Pam Samuelson demonstrated, courts have often found merger even when there was more than one way to express facts, as long as the number of ways was practically limited. See Samuelson, *supra* note 105, at 7-10.

but to identify the unprotectable elements of a work that still meets the threshold of protectability when considered as a whole. These elements are identified so that the substantial similarity analysis at the infringement stage can be focused on the protectable elements—although, as we describe further below, it’s not clear how well courts actually focus the inquiry this way. In *Leigh v. Warner Brothers, Inc.*, for example, the court held that Leigh’s copyright in his “photograph of a sculpture in the Bonaventure Cemetery known as the Bird Girl” extended only to “the selection of lighting, shading, timing, angle, and film.”¹⁰⁹ Leigh’s copyright did not cover “the appearance of the statue itself or of Bonaventure Cemetery, for Leigh ha[d] no rights in the statue or its setting.”¹¹⁰ Taking those unprotectable features into account, the court held that the district court was “correct to hold as a matter of law that the film sequences featuring the Bird Girl statue [in Warner Brothers’ film version of the novel] are not substantially similar to the protected elements of Leigh’s photograph,” but that there was a jury question regarding the similarity of Warner Brothers’ photographic images, which had “much in common with the elements protected by Leigh’s copyright.”¹¹¹

109. 212 F.3d 1210, 1213, 1215 (11th Cir. 2000). Leigh took the photo for the cover of *Midnight in the Garden of Good and Evil*, a novel by John Berendt. *Id.* at 1213 (“After reading a manuscript of the novel, Leigh explored appropriate settings in Savannah and ultimately selected a photograph of a sculpture in the Bonaventure Cemetery known as the Bird Girl. Sylvia Shaw Judson had sculpted the Bird Girl in 1938, and she produced three copies of the statue. The Trosdal family had purchased one of the statues and placed it in their plot at Bonaventure Cemetery. The novel does not mention the Bird Girl statue.”).

110. *Id.* at 1214. Courts sometimes suggest that “the merger doctrine is most applicable where the idea and the expression are of items found in nature, or are found commonly in everyday life.” *Yankee Candle Co. v. Bridgewater Candle Co.*, 259 F.3d 25, 36 (1st Cir. 2001).

111. *Leigh*, 212 F.3d at 1215-16.

Figure 1.
Bird Girl Statue (*Leigh v. Warner Brothers*)



We confess that we can't see much, if anything, in common between these two photos other than their core subject matter. But the court's approach treats idea-expression and related doctrines simply as inputs into the infringement analysis, and therefore as considerations for the jury at the infringement stage.¹¹²

Courts sometimes even treat merger as a scope doctrine. In *Yankee Candle Co. v. Bridgewater Candle Co.*, the court evaluated Yankee's claim that Bridgewater infringed the labels of nine candle fragrances by first dissecting the work to remove unprotectable elements and then "apply[ing] the doctrines of merger and scene-a-faire to determine *how* 'substantially similar' the copy must be to

112. The Ninth Circuit treats originality as a question of validity but treats merger as an affirmative defense to infringement, to be raised by the defendant. *See* *Apple Comput., Inc. v. Formula Int'l, Inc.*, 725 F.2d 521, 523 (9th Cir. 1984); *see also* *Oracle Am., Inc. v. Google Inc.*, 750 F.3d 1339, 1358 (Fed. Cir. 2014) (applying Ninth Circuit law). Nevertheless, the Ninth Circuit also maintains that merger is to be evaluated as of the time of creation of the work—meaning that the number of ways an idea can be expressed must be evaluated at the time the allegedly infringed work is created. *Oracle*, 750 F.3d at 1361. For a general discussion of the confused and circuit-dependent nature of dissection in infringement, see Mark A. Lemley, *Our Bizarre System for Proving Copyright Infringement*, 57 J. COPYRIGHT SOC'Y 719 (2010).

infringe.”¹¹³ Focusing on Yankee’s labels, the court found that, because the ideas of six of the nine labels merged with expression, “there could be no infringement unless Bridgewater’s photographs were ‘nearly identical’ to Yankee’s.”¹¹⁴ This was a matter of scope, however, and not validity. “The merger doctrine does not,” the court said, “allow the identical reproduction of photographs of realistic objects when there are sufficient details in those photographs to make them unique.”¹¹⁵ Thus:

If Bridgewater had scanned Yankee’s labels into a computer and reproduced them exactly, it would have certainly infringed Yankee’s copyrights on those labels. Even if Bridgewater had taken its own photographs, but had arranged the subjects in a “nearly identical” manner to that of Yankee, a jury could have found the requisite showing of substantial similarity to support copyright infringement.¹¹⁶

This kind of approach, which focuses very little on the validity stage and treats the limiting doctrines as inputs into the infringement analysis, puts tremendous pressure on courts to tailor the scope of rights in a work at the infringement stage. As courts typically recognize, the scope of protection to which an author is entitled is supposed to match the size of her original contribution.¹¹⁷ For that to work, validity and infringement have to be in alignment regarding which elements of a work deserve protection and the scope of protection to which a work is entitled.

In fact, existing infringement doctrine creates a number of potential gaps in the fabric of the scope inquiry. First, and most

113. *Yankee Candle*, 259 F.3d at 34 (citing *Concrete Mach. Co. v. Classic Lawn Ornaments, Inc.*, 843 F.2d 600, 609 n.9 (1st Cir. 1988)). On this approach, “the court may find that the idea and expression are so inseparable that copying of the work is not prohibited or that only exact reproduction of the work will justify a finding of infringement.” *Id.* (quoting *Concrete Mach.*, 843 F.2d at 609 n.9).

114. *Id.* at 35.

115. *Id.* at 36 (citing *Concrete Mach.*, 843 F.2d at 607-10).

116. *Id.*

117. See *Meshwerks, Inc. v. Toyota Motor Sales U.S.A., Inc.*, 528 F.3d 1258, 1264 (10th Cir. 2008) (“So, in the case of photographs, for which Meshwerks’ digital models were designed to serve as practically advantageous substitutes, authors are entitled to copyright protection only for the ‘incremental contribution,’ represented by their interpretation or expression of the objects of their attention.” (quoting *SHL Imaging, Inc. v. Artisan House, Inc.*, 117 F. Supp. 2d 301, 311 (S.D.N.Y. 2000))).

important, because of the way courts structure the infringement inquiry, there is no guarantee that the ordinary observer's determination of substantial similarity will be based on similarity only of protectable elements. Thus, although in theory copyright has a mechanism to identify the proper scope of protection, it often doesn't end up using that mechanism to decide cases.

There are two primary tests for infringement, both of which separate the question of actual copying from improper appropriation. Under the Second Circuit's approach in *Arnstein v. Porter*, to prove infringement, the plaintiff must establish "(a) that defendant copied from plaintiff's copyrighted work and (b) that the copying (assuming it to be proved) went so far as to constitute improper appropriation."¹¹⁸ On the first question of actual copying, "analysis ('dissection') is relevant, and the testimony of experts may be received to aid the trier of the facts. If evidence of access is absent, the similarities must be so striking as to preclude the possibility that plaintiff and defendant independently arrived at the same result."¹¹⁹ If actual copying is established, then the issue is "illicit copying"—what we call legal scope—and on that issue "the test is the response of the ordinary lay hearer," and dissection and expert testimony are irrelevant.¹²⁰

The other primary test is the Ninth Circuit's, which refers to the two steps as "extrinsic" and "intrinsic" comparisons.¹²¹ Under the "extrinsic" test, a court must determine whether there is substantial similarity in general ideas, and it should do so not on the basis of the responses of the trier of fact, but on "specific criteria which can be listed and analyzed," including the "type of artwork involved, the materials used, the subject matter, and the setting for the subject."¹²² As reformulated by the Ninth Circuit in *Shaw v. Lindheim* and *Cavalier v. Random House Co.*, this first step is not limited to weeding out ideas but extends to anything unprotectable that is

118. 154 F.2d 464, 468 (2d Cir. 1946).

119. *Id.*

120. *Id.* "In some cases," the court allowed, "the similarities between the plaintiff's and defendant's work are so extensive and striking as, without more, both to justify an inference of copying and to prove improper appropriation." *Id.* at 468-69.

121. *Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp.*, 562 F.2d 1157, 1165 (9th Cir. 1977).

122. *Id.* at 1164.

supposed to be filtered out.¹²³ Analytic dissection and expert testimony are relevant to the extrinsic analysis, and this type of similarity “may often be decided as a matter of law.”¹²⁴ If the works are similar under the extrinsic test, then the second question is whether there is substantial similarity of expression.¹²⁵ This similarity is to be assessed under the “intrinsic test,” which focuses on the impressions of the “average reasonable reader and spectator.”¹²⁶ On this issue, analytic dissection and expert testimony are inappropriate.¹²⁷ Substantial similarity under the intrinsic test is a question of fact; indeed, “satisfaction of the extrinsic test creates a triable issue of fact in a copyright action involving a literary work.”¹²⁸

These two formulations have at least one thing in common: both allow analytic dissection of a work to identify the protectable elements at the first stage, where the issue is actual copying, but reject dissection in the second stage where the question is whether the copying was improper.¹²⁹ As a result, notwithstanding the frequency with which courts say that the plaintiff must show substantial similarity between the defendant’s work and protectable elements of the plaintiff’s work, nothing in the doctrinal formulation actually requires that comparison. In fact, courts have found infringement in a variety of questionable cases in which the similarities seem largely to have related to unprotectable elements.¹³⁰

In *Copeland v. Bieber*, for instance, the Fourth Circuit reversed the district court’s rejection of Devin Copeland’s claim that Justin

123. See *Cavalier v. Random House, Inc.*, 297 F.3d 815 (9th Cir. 2002) (“[A] court must filter out and disregard the non-protectable elements in making its substantial similarity determination.”); *Shaw v. Lindheim*, 919 F.2d 1353 (9th Cir. 1990).

124. *Krofft*, 562 F.2d at 1164.

125. See *id.*

126. *Id.* (quoting *Twentieth Century-Fox Film Corp. v. Stonesifer*, 140 F.2d 579, 582 (9th Cir. 1944)).

127. See *id.*; see also *Copeland v. Bieber*, 789 F.3d 484, 489 (4th Cir. 2015) (“[A]nalytic dissection is inapplicable to the intrinsic analysis, because a work’s intended audience ‘does not make the distinction’ between protectable and unprotectable elements and instead encounters the work ‘as one object.’” (quoting *Universal Furniture Int’l, Inc. v. Collezione Europa USA, Inc.*, 618 F.3d 417, 437 (4th Cir. 2010))).

128. *Shaw*, 919 F.2d at 1359.

129. One of us has previously characterized this approach as “exactly backwards.” Lemley, *supra* note 112, at 719.

130. See, e.g., *Worlds of Wonder, Inc. v. Vector Intercont’l, Inc.*, No. C86-2671, 1986 WL 15608, at *1 (N.D. Ohio Dec. 31, 1986) (finding that a tape that controlled the Teddy Ruxpin doll’s movements infringed plaintiff’s copyright in its original doll).

Bieber had infringed his copyright in the song “I Need Somebody to Love.”¹³¹ In concluding that a jury might be able to find the songs similar enough for infringement despite numerous dissimilarities, the court focused on the fact that they both include the same title lyric, “I need somebody to love.”¹³² Bieber pointed out that the lyric was not original to Copeland, but the court was having none of it.¹³³ Identifying the copied element as unprotectable was analytic dissection, and analytic dissection was not appropriate in the intrinsic test, which focused on look and feel.¹³⁴ Whatever the similarities between the songs, this is a ridiculous conclusion on its face; it would allow a copyright claim against someone who explicitly copied only things in the public domain over which the copyright owner had no claim. But the Fourth Circuit is hardly alone in this artificial distinction.¹³⁵

Courts have sometimes recognized the problem and tried to ameliorate it by departing from the *Arnstein* approach and allowing dissection, and in some cases even expert testimony, at the second step. As the Seventh Circuit said, “[w]hile dissection is generally disfavored, the ordinary observer test, in application, must take into account that the copyright laws preclude appropriation of only those elements of the work that are protected by the copyright.”¹³⁶ Other courts have tried to focus the ordinary observer inquiry by having the court first identify which elements of the plaintiff’s work are protectable. In *Yankee Candle*, for example, the First Circuit said that, for purposes of evaluating substantial similarity:

131. See 789 F.3d at 495.

132. See *id.* at 494.

133. See *id.* at 492-93.

134. See *id.* at 489.

135. Among other cases, the Ninth Circuit in *Meridian Textiles, Inc. v. Topson Downs of California, Inc.* held that a jury was entitled to decide whether the defendant’s use of a zebra stripe design infringed the plaintiff’s zebra stripe design. See 605 F. App’x 671, 672 (9th Cir. 2015). The district court had granted summary judgment on the grounds that the plaintiff’s design was too similar to every other zebra stripe out there to be protectable, but the court reversed, saying that the plaintiff could have a copyright because there was no evidence it copied from the prior art. See *id.* In a separate part of the opinion, the court then concluded that the jury would get to decide the similarity of the two designs, apparently without the benefit of the fact that there was nothing in the plaintiff’s design that was not in countless prior designs. See *id.*

136. *Atari, Inc., v. N. Am. Phillips Consumer Elecs. Corp.*, 672 F.2d 607, 614 (7th Cir. 1982); see also *Susan Wakeen Doll Co. v. Ashton-Drake Galleries*, 272 F.3d 441, 451 (7th Cir. 2001) (noting, approvingly, that the jury had heard expert testimony regarding similarities).

The extent to which [plaintiff's work] contain[s] protected expression is a matter of law, determined by the court. Once this determination is made, the question of whether two works are substantially similar (and corresponding application of the ordinary observer test) is a matter for the trier of fact unless summary judgment is proper.¹³⁷

Similarly, in *Apple Computer, Inc. v. Microsoft Corp.*, the Ninth Circuit held that in a case involving computer software, once the plaintiff had satisfied the extrinsic test, a court should frame the comparison for purposes of the intrinsic test.¹³⁸ "Using analytic dissection, and, if necessary, expert testimony, the court must determine whether any of the allegedly similar features are protected by copyright."¹³⁹ Then,

[h]aving dissected the alleged similarities and considered the range of possible expression, the court must define the scope of the plaintiff's copyright—that is, decide whether the work is entitled to "broad" or "thin" protection. Depending on the degree of protection, the court must set the appropriate standard for a subjective comparison of the works to determine whether, as a whole, they are sufficiently similar to support a finding of illicit copying.¹⁴⁰

In a later case, the Ninth Circuit emphasized that, at least in software cases, "[a]nalytic dissection is relevant not only to the copying element of a copyright infringement claim, but also to the claim's ownership element."¹⁴¹

Courts also sometimes ratchet up the level of similarity necessary to find infringement in cases in which copyright protection is supposed to be "thin." In these cases, courts have required that the

137. *Yankee Candle Co. v. Bridgewater Candle Co.*, 259 F.3d 25, 34 n.5 (1st Cir. 2001) (citing *Concrete Mach. Co. v. Classic Lawn Ornaments, Inc.*, 843 F.2d 600, 608-09 (1st Cir. 1988)).

138. 35 F.3d 1435, 1443 (9th Cir. 1994).

139. *Id.* at 1443.

140. *Id.*; see also *Brown Bag Software v. Symantec Corp.*, 960 F.2d 1465, 1475 (9th Cir. 1992) (approving of the district court's dissection of the plaintiff's work for the "purpose of determining 'whether similarities [between the programs] result from unprotectable [or protected] expression'" (quoting *Data E. USA, Inc. v. Epyx, Inc.*, 862 F.2d 204, 208 (9th Cir. 1988))) (alteration in original).

141. *Brown Bag*, 960 F.2d at 1476.

defendant's work be not just substantially similar to the plaintiff's, but "virtually identical,"¹⁴² or they have required substantial similarity from the perspective of the "more discerning observer," rather than the "ordinary observer."¹⁴³

Courts seem most eager to depart from the two-step process and allow dissection and expert testimony in cases involving software, reflecting concern that ordinary observers will have too much difficulty assessing similarity of such complex works.¹⁴⁴ Several courts have followed the Second Circuit's abstraction-filtration-comparison approach from *Computer Associates International, Inc. v. Altai, Inc.*, a particularly structured approach to substantial similarity under which a court first dissects the plaintiff's work into elements at varying levels of abstraction, then filters out the unprotectable elements, and finally compares the allegedly infringing work to the remaining "kernels" of creative expression.¹⁴⁵ But this approach seems to treat software as a special case. Elsewhere the notion that particular copyrightable elements should be flagged for purposes of comparison is in constant tension with the requirement that the works be compared "as a whole," particularly in those cases that emphasize the "total look and feel" of the works.¹⁴⁶

Lack of clarity about whether and when the substantial similarity analysis will explicitly take scope into account is exacerbated by the fact that allocation of decisionmaking in copyright cases is

142. See, e.g., *Apple*, 35 F.3d at 1442; *Hoehling v. Universal City Studios, Inc.*, 618 F.2d 972, 979-80 (2d Cir. 1980).

143. See, e.g., *Boisson v. Banian, Ltd.*, 273 F.3d 262, 271-72 (2d Cir. 2001).

144. See, e.g., *Whelan Assocs., Inc. v. Jaslow Dental Lab., Inc.*, 797 F.2d 1222, 1233 (3d Cir. 1986) ("We therefore join the growing number of courts which do not apply the ordinary observer test in copyright cases involving exceptionally difficult materials, like computer programs, but instead adopt a single substantial similarity inquiry according to which both lay and expert testimony would be admissible.").

145. 982 F.2d 693, 706 (2d Cir. 1992); see also *Eng'g Dynamics, Inc. v. Structural Software, Inc.*, 26 F.3d 1335, 1342-43 (5th Cir. 1994); *Gates Rubber Co. v. Bando Chem. Indus., Ltd.*, 9 F.3d 823, 834 (10th Cir. 1993). On the consensus behind the *Altai* framework, see Mark A. Lemley, *Convergence in the Law of Software Copyright?*, 10 HIGH TECH. L.J. 1, 14-15 (1995).

146. Courts seem particularly drawn to the "total feel" approach in cases involving visual works. See, e.g., *Tufenkian Imp./Exp. Ventures, Inc. v. Einstein Moomjy, Inc.*, 338 F.3d 127, 134 (2d Cir. 2003) ("Essentially, the total-concept-and-feel locution functions as a reminder that, while the infringement analysis must *begin* by dissecting the copyrighted work into its component parts in order to clarify precisely what is not original, infringement analysis is not *simply* a matter of ascertaining similarity between components viewed in isolation."); *Roth Greeting Cards v. United Card Co.*, 429 F.2d 1106, 1110 (9th Cir. 1970).

remarkably unclear. Some courts treat copyrightability as a question of fact for purposes of assessing validity.¹⁴⁷ The Seventh Circuit treats it as a question of law.¹⁴⁸ The Ninth Circuit has issued conflicting decisions on the respective roles of judge and jury.¹⁴⁹ When validity and infringement are determined by different decision-makers, there is no reason to be confident that the scope limitations presumed at the validity stage will in fact be reflected at the infringement stage.

In the now-famous “Blurred Lines” copyright trial, for instance, the jury was asked to decide only whether the written chords and lyrics of Robin Thicke and Pharrell Williams’s hit song were too similar to Marvin Gaye’s song “Got to Give It Up.”¹⁵⁰ Because Marvin Gaye did not have a copyright that extended to his tempo or vocal style, the relevant comparison should have been between Thicke and Williams’s song and Gaye’s written work, not between the sound of the two songs.¹⁵¹ But it is hard to believe the jury actually ignored similarities in the unprotectable elements, or the evidence of the defendants’ intent to imitate Gaye’s style. And if they did in fact do as they were told, it is hard to see how they could have found sufficient similarity. By handing the overall works to the jury, the court undid whatever limits the rules of copyright placed on the scope of Gaye’s right.¹⁵²

147. See, e.g., *Matthew Bender & Co. v. W. Publ’g Co.*, 158 F.3d 674, 681 (2d Cir. 1998) (“[W]e treat the question of whether particular elements of a work demonstrate sufficient originality and creativity to warrant copyright protection as a question for the factfinder.”).

148. See *Schrock v. Learning Curve Int’l, Inc.*, 586 F.3d 513, 517 (7th Cir. 2009) (“In this circuit, copyrightability is an issue of law for the court.”).

149. Compare *Ets-Hokin v. Skyy Spirits, Inc.*, 225 F.3d 1068, 1073 (9th Cir. 2000) (“Whether a particular photograph is protected by copyright law is a mixed question of law and fact, also subject to de novo review.”), with *Meridian Textiles, Inc. v. Topson Downs of Cal., Inc.*, 605 F. App’x 671, 672 (9th Cir. 2015) (treating the originality of the plaintiff’s work as a question of fact for the jury), and *N. Coast Indus. v. Jason Maxwell, Inc.*, 972 F.2d 1031, 1034 (9th Cir. 1992) (same).

150. See *Williams v. Bridgeport Music, Inc.*, No. LA CV13-06004 JAK (AGRx), 2015 WL 4479500, at *1 (C.D. Cal. July 14, 2015).

151. See *id.* at *18.

152. See Wendy Gordon, *How the Jury in the ‘Blurred Lines’ Case Was Mised*, CONVERSATION (Mar. 17, 2015, 5:47 AM), <http://theconversation.com/how-the-jury-in-the-blurred-lines-case-was-mised-38751> [<https://perma.cc/C6YK-LJP9>] (questioning the instructions given to the jury); Chris Richards, *It’s Okay if You Hate Robin Thicke. But the ‘Blurred Lines’ Verdict Is Bad for Pop Music*, WASH. POST (Mar. 11, 2015), <http://www.washingtonpost.com/news/arts-and-entertainment/wp/2015/03/11/the-blurred-lines-of-the-blurred-lines-verdict/> [<https://perma.cc/9RWG-925S>] (noting the dissimilarity between the lyrics and notes).

Misalignment may arise even within the infringement analysis because allocation of decisionmaking between the two parts of the tests is also unclear. According to *Arnstein*, both actual copying and improper appropriation are issues of fact.¹⁵³ The *Yankee Candle* and *Apple v. Microsoft* approaches maintain the factfinder's role in making the ultimate assessment of substantial similarity, but they require the court to frame the comparison at the improper appropriation (or intrinsic comparison) stage, implying either that dissection is for the court or that improper appropriation at least requires guidance from the court.¹⁵⁴

The problems with the fracturing of these inquiries is worse when different decisionmakers are involved at the various stages, but we think the prospect for misalignment is real even when the same decisionmaker is responsible for all of the determinations. Indeed, recent empirical work by Shyamkrishna Balganes, Irina Manta, and Tess Wilkinson-Ryan suggests that judgments of similarity may often be contaminated by information relating to the defendant's actual copying.¹⁵⁵ Specifically, the authors find "an appreciable upward shift in subjects' assessments of similarity between the works" when they were "provided with additional information about the simple fact of copying or the creative effort that went into the protected work."¹⁵⁶ And if evidence of copying affects even the judgment of similarity, it's hard to believe that a decisionmaker who has already concluded that the defendant copied could focus the substantial similarity analysis on particular protectable elements

153. See *Arnstein v. Porter*, 154 F.2d 464, 469 (2d Cir. 1946); see also *Leigh v. Warner Bros., Inc.*, 212 F.3d 1210, 1213 (11th Cir. 2000) ("Copyright infringement is generally a question of fact for the jury to decide."). For an argument that *Arnstein* was in fact about the fact-law distinction and the proper role of judge and jury, and only secondarily about the standard for copyright infringement, see Shyamkrishna Balganes, *The Questionable Origins of the Copyright Infringement Analysis* 6 (Univ. of Pa. Law Sch. Pub. Law & Legal Theory Research Paper Series, Paper No. 15-28, 2015), <http://ssrn.com/abstract=2654537> [<https://perma.cc/DP8E-G3NH>].

154. See *supra* notes 113-17 and 139-41 and accompanying text.

155. See Shyamkrishna Balganes et al., *Judging Similarity*, 100 IOWA L. REV. 267, 270-71 (2014).

156. *Id.* at 271. The study's methodology assumed that the party assessing similarity would have encountered information about actual copying but would not actually have had to arrive at that conclusion herself. That arrangement obviously does not match the decisionmaking architecture in all jurisdictions. It seems reasonable, however, to conclude that the contamination effect might even be stronger when a particular decisionmaker herself determined that the defendant had actually copied.

rather than considering evidence of copying of unprotectable pieces as well.¹⁵⁷

While the errors in copyright largely run in the direction of over-protection because of insufficient focus on the protectable elements at the substantial similarity phase, sometimes that problem causes errors in the other direction. Courts have, for example, sometimes found merger despite the fact that there seemed to be many other ways to express the idea underlying the plaintiff's work. They have done so because of uncertainty about how to determine which of those alternatives would infringe, typically where the plaintiff overclaimed. In *Herbert Rosenthal Jewelry Corp. v. Kalpakian*, for example, the court held that the plaintiff could not claim copyright in a jeweled bee pin because allowing protection would give it a monopoly on depicting bees in jeweled pins.¹⁵⁸

Figure 2.

Jeweled Bee Pin (*Herbert Rosenthal Jewelry Corp. v. Kalpakian*)



157. In *Whelan*, the court suggested that if the same decisionmaker was involved at both stages of the substantial similarity inquiry, it might not matter whether expert testimony and dissection are technically relevant at the improper appropriation stage. *Whelan Assocs., Inc. v. Jaslow Dental Lab., Inc.*, 797 F.2d 1222, 1232-33 (3d Cir. 1986) (“[T]he distinction between the two parts of the *Arnstein* test may be of doubtful value when the finder of fact is the same person for each step: that person has been exposed to expert evidence in the first step, yet she or he is supposed to ignore or ‘forget’ that evidence in analyzing the problem under the second step. Especially in complex cases, we doubt that the ‘forgetting’ can be effective when the expert testimony is essential to even the most fundamental understanding of the objects in question.”). The Balganesch study suggests the court was right to doubt the decisionmaker’s ability to “forget,” but that the problem of forgetting is more likely to relate to evidence of actual copying. See Balganesch et al., *supra* note 155, at 269.

158. See 446 F.2d 738, 742 (9th Cir. 1971).

That result has struck some observers as bizarre, because there are obviously many ways to design a jeweled bee pin.¹⁵⁹ But we think that criticism misses the obvious point that the plaintiff asked the court to recognize rights of astonishing scope. As the court said:

The breadth of this claim is evident. For example, while a photograph of the copyrighted bee pin attached to the complaint depicts a bee with nineteen small white jewels on its back, plaintiff argues that its copyright is infringed by defendants' entire line of a score or more jeweled bees in three sizes decorated with from nine to thirty jewels of various sizes, kinds, and colors.

Although plaintiff's counsel asserted that the originality of plaintiff's bee pin lay in a particular arrangement of jewels on the top of the pin, the elements of this arrangement were never identified. Defendants' witnesses testified that the "arrangement" was simply a function of the size and form of the bee pin and the size of the jewels used. Plaintiff's counsel, repeatedly pressed by the district judge, was unable to suggest how jewels might be placed on the back of a pin in the shape of a bee without infringing plaintiff's copyright. He eventually conceded, "not being a jeweler, I can't conceive of how he might rearrange the design so it is dissimilar."¹⁶⁰

As the court rightly recognized, "[i]f plaintiff's understanding of its rights were correct, its copyright would effectively prevent others from engaging in the business of manufacturing and selling jeweled bees."¹⁶¹ In an ideal world, the court would have had at its disposal a mechanism by which the protectable elements of the plaintiff's bee pins could have been identified, and the court then could have resolved the particular case by comparing the defendant's bee pins to those protectable elements. In the absence of a reliable mechanism of that sort, the court felt it had no choice but to declare the bee pins uncopyrightable.¹⁶²

159. To wit, a search of "bee pin" on Amazon.com, limited to jewelry, returns nearly 500 results from more than 50 designers.

160. *Herbert Rosenthal Jewelry Corp.*, 446 F.2d at 740.

161. *Id.*

162. *See id.* at 742.

2. Patent

As we will see in the next Part, utility patent law does more than its sister doctrines to evaluate the scope of a patent right in an integrated fashion. The patent claims themselves are an effort to define the scope of the legal right *ex ante*. And the pretrial *Markman* hearing requires the parties to resolve the scope of the patent right before either validity or infringement is established.¹⁶³

Nonetheless, patent law too has gaps that result from its separation of validity and infringement. The first and most obvious of these gaps involves defendants who practice a technology that is already in the prior art. Because patent law requires that an invention be novel and nonobvious,¹⁶⁴ it should be clear that a patent owner is not entitled to sue someone for using technology that existed before she ever “invented” it. Nonetheless, the Federal Circuit has repeatedly held that there is no “practicing the prior art” defense in patent law.¹⁶⁵ As Timothy Lau has ably explained, this rather bizarre rule results from the separation of patent law doctrines into validity and infringement.¹⁶⁶ Because the patent owner bears the burden of proving infringement but the defendant must prove invalidity by clear and convincing evidence, courts worry that defendants can skimp on their burden of proof that the defendant’s technology was in fact in the prior art.¹⁶⁷ But the fact that the courts will not recognize practicing the prior art as a defense means that defendants who are in fact engaged in what should be permitted activity do not always have a fair opportunity to prove it. In theory, the *Markman* hearing solves this problem by treating validity and infringement together in an integrated proceeding.¹⁶⁸ But the

163. See *supra* note 9 and accompanying text. In part for this reason, Janson Rantanen has argued that patent rights, which seem to be immutable once the claims are written, are in fact more malleable and contextual than one would think. See Jason Rantanen, *The Malleability of Patent Rights*, 2015 MICH. ST. L. REV. 895, 898.

164. See 35 U.S.C. §§ 102-103 (2012).

165. See, e.g., *Kinetic Concepts, Inc. v. Blue Sky Med. Grp., Inc.*, 554 F.3d 1010, 1025 (Fed. Cir. 2009) (“[P]racticing the prior art’ is not a defense to patent infringement.”); *Tate Access Floors, Inc. v. Interface Architectural Res., Inc.*, 279 F.3d 1357, 1365 (Fed. Cir. 2002) (“[T]here is no ‘practicing the prior art’ defense to literal infringement.”).

166. See Timothy Lau, *Defensive Use of Prior Art to Exonerate Accused Acts in U.S. and Chinese Patent Litigation*, 27 COLUM. J. ASIAN L. 51, 60-61 (2013).

167. See Lemley, *supra* note 24, at 1271, 1276.

168. See *supra* note 9 and accompanying text.

Federal Circuit has discouraged courts from focusing much attention on validity during *Markman*.¹⁶⁹ A district court may construe the patent claims for infringement purposes without focusing on the possibility that a claim construed so broadly might be invalid. And when the issue of validity does arise, district courts and juries may be reluctant (even apart from the high standard of proof) to hold the patent altogether invalid.

Patent owners can and do exploit this gap with some regularity. Relying both on the higher standard of proof for invalidity than non-infringement and the demonstrated reluctance of juries to invalidate patents, patent owners are often willing to trade a greater risk of eventual invalidity to gain a broad scope for their possibly invalid patent claims.¹⁷⁰ Sometimes this gamble fails. In *Liebel-Flarsheim Co. v. Medrad, Inc.*, for instance, the patentee succeeded in persuading the court to construe its claim to be much broader than the thing it had actually invented,¹⁷¹ only to have the claim ultimately invalidated as insufficiently supported by the patent disclosure.¹⁷² But more often the strategy succeeds, either in court or in settlement before trial.¹⁷³ This is a particular problem with “bottom-feeder”

169. While it was traditionally a rule of patent claim construction that claims were to be construed to preserve their validity, that canon is in disrepute in the Federal Circuit right now. See *Phillips v. AWH Corp.*, 415 F.3d 1303, 1327-28 (Fed. Cir. 2005) (en banc) (dismissing that canon as one “of limited utility in any event” and saying that “we have not applied that principle broadly”). The reason once again can be traced to a scope gap: different decision-makers evaluate claim construction and infringement, and they do so at different times. The Federal Circuit was concerned about treading on the province of the jury to evaluate validity. See *id.* at 1330 (Mayer, J., dissenting). But the result has been that claim construction and infringement are assessed without a clear idea of whether the patentee is seeking excessive scope. See Timothy R. Holbrook, *Possession in Patent Law*, 59 SMU L. REV. 123, 160 (2006).

170. See John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185, 212 (1998) (finding that juries uphold patents 67 percent of the time, while judges in bench trials do so only 57 percent of the time). More recent evidence suggests that the gap has disappeared, though that trend seems to be driven primarily by Abbreviated New Drug Application (ANDA) cases, which are decided by judges, not juries. See John R. Allison et al., *Understanding the Realities of Modern Patent Litigation*, 92 TEX. L. REV. 1769, 1790 (2014); Lemley, *supra* note 72, at 1734.

171. 358 F.3d 898, 912 (Fed. Cir. 2004).

172. *Liebel-Flarsheim Co. v. Medrad, Inc.*, 481 F.3d 1371, 1383 (Fed. Cir. 2007). For a discussion of other cases presenting this sort of overclaiming, see Holbrook, *supra* note 7, at 800-03.

173. The overwhelming majority of patent cases settle before trial. See, e.g., John R. Allison et al., *Patent Quality and Settlement Among Repeat Patent Litigants*, 99 GEO. L.J. 677, 689 (2011) (finding that between 84 and 91 percent of cases settle); Jay P. Kesan & Gwendolyn G. Ball, *How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and*

patent trolls who are only interested in coercing a nuisance-value settlement based on the cost and uncertainty of litigation.¹⁷⁴ The practical effect of not having a practicing the prior art defense is that some defendants who are actually using old technology are nonetheless held liable, and more are forced to pay settlements to avoid the risk of losing a case that, in theory, they should win.

A final example of a scope gap in patent law operates against patentees rather than for them. Patent law is designed to cover new inventions but not to preempt abstract ideas altogether.¹⁷⁵ Patentees in computer software, following the playbook that has worked for them in the practicing the prior art example, have sought broader and broader interpretations of their patent claims, to the point where many claims are not limited either to a particular computer algorithm or approach or to a particular hardware implementation. Rather, they claim any computer configured in any way to achieve a particular result.¹⁷⁶ But in this case, the backlash against that functional claiming has been spectacular. The Supreme Court held in 2014 that claims of this sort are not patent-eligible at all,¹⁷⁷ and the lower courts have already invalidated dozens of patents on this ground.¹⁷⁸ The fact that the scope determination got pushed into

Settlement of Patent Disputes, 84 WASH. U. L. REV. 237, 273-74 (2006) (finding that between 65 and 68 percent of cases settle).

174. For a discussion of the bottom-feeder troll business model, see Mark A. Lemley & A. Douglas Melamed, *Missing the Forest for the Trolls*, 113 COLUM. L. REV. 2117, 2121-24 (2013).

175. See, e.g., *Alice Corp. Pty. Ltd. v. CLS Bank Int'l*, 134 S. Ct. 2347, 2354-55 (2014) (“[W]e must distinguish between patents that claim the ‘buildin[g] block[s]’ of human ingenuity and those that integrate the building blocks into something more, thereby ‘transform[ing]’ them into a patent-eligible invention. The former ‘would risk disproportionately tying up the use of the underlying’ ideas and are therefore ineligible for patent protection. The latter pose no comparable risk of preemption, and therefore remain eligible for the monopoly granted under our patent laws.”) (alterations in original) (citations omitted); *Bilski v. Kappos*, 561 U.S. 593, 609 (2010) (rejecting patent claim on the grounds that “it claims an abstract idea”).

176. See Kevin Emerson Collins, *Patent Law’s Functionality Malfunction and the Problem of Overbreadth, Functional Software Patents*, 90 WASH. U. L. REV. 1399, 1405 (2013) (discussing the overbreadth problem and advocating an algorithmic approach to analysis); Mark A. Lemley, *Software Patents and the Return of Functional Claiming*, 2013 WIS. L. REV. 905, 907 (discussing the breadth of this phenomenon and the problems it causes); Kyle O. Logan, *Step-Plus-Function Claims: An Analysis of Federal Circuit Law*, 24 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 907 (2014) (looking at the issue of overbreadth as a failure of the construction of 35 U.S.C. § 112(f)).

177. See *Alice*, 134 S. Ct. at 2360.

178. See, e.g., *Ultramercial, Inc. v. Hulu, LLC*, 772 F.3d 709, 722 (Fed. Cir. 2014) (Mayer, J., concurring) (rejecting patent claim on the “concept ... that people will be willing to watch

validity made it an all-or-nothing exercise, and in this case patentees end up getting nothing as a result of their overreaching, even if they would have a valid narrower claim to a particular software algorithm.

3. Trademark

Trademark law has, especially relative to copyright, several validity doctrines with some bite. Most importantly for our purposes, valid marks must be distinctive and nonfunctional.¹⁷⁹ But courts tend to rely primarily on the likelihood of confusion factors and on certain defendant-specific doctrines like descriptive fair use to manage the scope of trademark rights. Validity and infringement are interrelated here. Distinctiveness, trademark law's core validity concept, significantly affects scope through the likelihood of confusion factors precisely because distinctiveness can only be assessed in relation to particular goods or services.¹⁸⁰ Thus, though trademark rights are no longer strictly limited to the goods for which the plaintiff uses the mark,¹⁸¹ those goods continue to anchor the likelihood of confusion analysis through the similarity of goods factor.¹⁸²

online advertisements in exchange for the opportunity to view copyrighted material"); *buySAFE, Inc. v. Google, Inc.*, 765 F.3d 1350, 1355 (Fed. Cir. 2014) (rejecting a patent claim related to online transactions as "squarely about creating a contractual relationship" and thus "directed to an abstract idea"); *Digitech Image Techs., LLC v. Elecs. for Imaging, Inc.*, 758 F.3d 1344, 1351 (Fed. Cir. 2014) (applying the *Alice* test to patent claim related to digital image processing to reject the claim as directed to an abstract idea).

179. *See* 15 U.S.C. § 1052(d), (e)(5) (2012).

180. Courts determine the distinctiveness of word marks by placing them in categories along the *Abercrombie* spectrum according to the relationship between the mark and the goods. *See Abercrombie & Fitch Co. v. Hunting World, Inc.*, 537 F.2d 4, 9 (2d Cir. 1976).

181. *See Lemley & McKenna, supra* note 67, at 140 (highlighting a trademark claim by Soccer United Marketing against Black & Decker as an example of "trademark owners ... asserting the right to own markets themselves because ... the relevant market owes its origin to their brands").

182. The well-known *Polaroid* factors were developed explicitly to gauge the likelihood of confusion in the case of noncompeting products. *See Polaroid Corp. v. Polarad Elecs. Corp.*, 287 F.2d 492, 495 (2d Cir. 1961) ("Where the products are different, the prior owner's chance of success is a function of many variables: the strength of his mark, the degree of similarity between the two marks, the proximity of the products, the likelihood that the prior owner will bridge the gap, actual confusion, and the reciprocal of defendant's good faith in adopting its own mark, the quality of defendant's product, and the sophistication of the buyers."). Each of the courts of appeals has its own, nonexhaustive set of likelihood of confusion factors, though there is substantial overlap among them. *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION

But similarity of goods is only one dimension of relevant similarity, and infringement is determined by evaluating the confusion factors flexibly, with no one factor being dispositive.¹⁸³ In general, we can say that the more similar the mark and the more similar the goods or services, the more likely the use infringes, but there is no definitive threshold level of similarity of mark or of goods that makes another's use noninfringing.¹⁸⁴ And, importantly from our perspective, while the confusion factors require assessment of confusion as a result of the defendant's use of a mark and courts therefore often focus on features of the relevant marks, none of the factors explicitly seek to ascertain whether the features of the defendant's mark that are causing the confusion are features that define the plaintiff's protectable interest.

This isn't to say that the confusion analysis doesn't in some ways take scope into account. In fact, it is black letter law that the amount of similarity necessary to support a likelihood of confusion finding is a function of the strength of the plaintiff's trademark.¹⁸⁵ Strong trademarks get broader protection.¹⁸⁶ And while strength

§§ 20-23 (AM. LAW INST. 1995).

183. See *Nora Beverages, Inc. v. Perrier Grp. of Am., Inc.*, 269 F.3d 114, 119 (2d Cir. 2001) (noting that although the *Polaroid* factors are important in determining infringement, no one factor is dispositive and courts should focus on "the ultimate question of whether consumers are likely to be confused" (quoting *Paddington Corp. v. Altiki Imps. & Distribs., Inc.*, 996 F.2d 577, 584 (2d Cir. 1993))). *But cf.* *Barton Beebe, An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581, 1587 (2006) ("A finding that the similarity of the marks factor does not favor a likelihood of confusion is, in practice, dispositive, and a finding that the proximity of the goods factor does not favor a likelihood of confusion is nearly dispositive.... Meanwhile, ... a finding of bad faith intent is in fact dispositive.").

184. See *Jada Toys, Inc. v. Mattel, Inc.*, 518 F.3d 628, 632 (9th Cir. 2007) ("[W]e have never countenanced a likelihood of confusion determination based on a consideration of dissimilarity alone.").

185. See *Virgin Enters. Ltd. v. Nawab*, 335 F.3d 141, 147 (2d Cir. 2003) ("The strength of a trademark encompasses two different concepts, both of which relate significantly to likelihood of consumer confusion. The first and most important is inherent strength, also called 'inherent distinctiveness.' This inquiry distinguishes between, on the one hand, inherently distinctive marks—marks that are arbitrary or fanciful in relation to the products (or services) on which they are used—and, on the other hand, marks that are generic, descriptive or suggestive as to those goods. The former are the strong marks. The second sense of the concept of strength of a mark is 'acquired distinctiveness,' *i.e.*, fame, or the extent to which prominent use of the mark in commerce has resulted in a high degree of consumer recognition.") (citation omitted).

186. See *id.* ("[T]he law accords broad, muscular protection to marks that are arbitrary or fanciful in relation to the products on which they are used, and lesser protection, or no protection at all, to marks consisting of words that identify or describe the goods or their

relates primarily to consumer recognition of a mark, in assessing strength many courts consider whether others in the relevant market also use the claimed mark to identify their own goods or services.¹⁸⁷ When a number of parties are using the same or a similar mark, each party's rights are narrow.¹⁸⁸ In *Eastland Music Group, LLC v. Lionsgate Entertainment, Inc.*, for example, the court noted the extensive prior use of "the phrase 50/50 or a sound-alike variant (50-50, fifty-fifty, fifty/fifty)" as the "title of intellectual property" and therefore held that, by entering such a crowded field, Eastland's rights in its Phifty-50 mark, which it used for a rap duo, were "weak and narrow."¹⁸⁹ The Federal Circuit recently expounded on this doctrine at length, suggesting that it may play a greater role at the PTO at least.¹⁹⁰

Courts have been most successful in aligning scope at the validity and infringement stages when, as in *Eastland*, they expressly

attributes.").

187. See *First Sav. Bank, F.S.B. v. First Bank Sys., Inc.*, 101 F.3d 645, 654 (10th Cir. 1996) ("[W]e recognized the well-established principle that extensive third-party use of the disputed term indicates that the term itself deserves only weak protection."). When assessing strength in the likelihood of confusion contexts, courts have sometimes looked more broadly, viewing as relevant the extent to which a mark has been used by others in any market. See, e.g., *Armco, Inc. v. Armco Burglar Alarm Co.*, 693 F.2d 1155, 1159 (5th Cir. 1982) (noting that absence of third-party use contributes to the strength of a trademark); *Sun Banks of Fla., Inc. v. Sun Fed. Sav. & Loan Ass'n*, 651 F.2d 311, 316 (5th Cir. 1981) ("[W]e find the extensive third-party use of the word 'Sun' impressive evidence that there would be no likelihood of confusion between Sun Banks and Sun Federal."); *Amstar Corp. v. Domino's Pizza, Inc.*, 615 F.2d 252, 259-60 (5th Cir. 1980) (finding extensive third-party uses of "Domino" limited protection of plaintiff's mark outside the uses to which it had already put its mark). Sara Stadler Nelson has thus argued that uniqueness is really what the courts are interested in when they talk about strength. See Sara Stadler Nelson, *The Wages of Ubiquity in Trademark Law*, 88 IOWA L. REV. 731, 734-36 (2003).

188. See 2 MCCARTHY, *supra* note 23, § 11:85 ("In a 'crowded' field of look-alike marks, each member of the crowd is relatively 'weak' in its ability to prevent use by others in the crowd.").

189. 707 F.3d 869, 871 (7th Cir. 2013). To our point about the difficulty of separating issues between validity and infringement, the Seventh Circuit also noted that "the principal reason [Phifty-50] was registrable is that it is a made-up homophone of a familiar phrase, which in ordinary usage is suggestive or descriptive." *Id.* Implying then that perhaps the issue was actually distinctiveness, the court said "[i]t takes a powerful showing of association between such an expression and a particular producer of goods to establish a trademark claim—and Eastland Music has not attempted such a showing." *Id.*

190. See *Juice Generation, Inc. v. GS Enters. LLC*, 794 F.3d 1334, 1338-39 (Fed. Cir. 2015) (holding that the Peace and Love trademark, while valid, was narrow because it was diluted by the presence of many other similar marks, and therefore could not preclude defendant from selling "Peace Love and Juice" drinks).

consider those two issues together and specifically describe the aspects of the plaintiff's mark that warrant protection. In *Star Industries, Inc. v. Bacardi & Co.*, the court found Star Industries's "O" design inherently distinctive because it "had sufficient shape and color stylization to render it slightly more than a simply linear representation of an ellipse or the letter 'O'" and was "a unique design in the alcoholic beverage industry at the time it was introduced."¹⁹¹ But, the court noted, "the extent of stylization was marginal at best. The outline of the 'O,' though not uniform, is ordinary in its slightly varying width, and the interior and exterior borders are also ordinary."¹⁹² As a result, Star Industries had only a "thin' or weak mark, which [was] entitled to only limited protection."¹⁹³ Having called out the specific aspects of Star's "O" design that made it protectable, the court could compare Bacardi's "O" with those specific features in mind. Unsurprisingly, the court then found that Bacardi's orange "O" design was not likely to cause confusion.

Figure 3.
"O" Design (*Star Industries v. Bacardi*)



In our view, a court should come to a similar conclusion in Adidas's claim against Marc Jacobs. Adidas, which owns a three-stripe design mark for shoes that it also applies to clothing, has sued Marc Jacobs for putting four stripes on its clothes.¹⁹⁴ Despite

191. 412 F.3d 373, 383 (2d Cir. 2005).

192. *Id.*

193. *Id.*

194. Complaint, *Adidas Am., Inc. v. Marc Jacobs Int'l, LLC*, No. 03:15-cv-582, 2015 WL 1569601 (D. Or. Apr. 8, 2015); see Jonathan Randles, *Adidas Accuses Marc Jacobs of Ripping Off 3-Stripe Design*, LAW360 (Apr. 8, 2015, 6:15 PM), <http://www.law360.com/articles/640950/adidas-accuses-marc-jacobs-of-ripping-off-3-stripe-design> [<https://perma.cc/VTJ2-QVFK>].

the similarity in some of the colors and the width, we doubt that a mark consisting of three vertical stripes could be broad enough in scope to cover anything but an essentially exact copy.¹⁹⁵

Figure 4.
Three-Stripe Mark (*Adidas v. Marc Jacobs*)



But courts have had a much easier time doing this with relatively simple marks—single letter designs like a stylized “O” or combinations of letters—than with compound, contextual marks. Thus, we can imagine a court applying this principle to deal with Lagunitas Brewing’s recent suit against Sierra Nevada Brewing.¹⁹⁶ Lagunitas claimed that Sierra Nevada infringed its rights in an “‘IPA’ family of trademarks.”¹⁹⁷ Lagunitas admitted that “other brewers have adopted the shorthand parlance of ‘IPA’ to market their India Pale Ales,” and indeed the “IPA” designation is in widespread use by a

195. For one thing, unlike when they are used on the side of athletic shoes, we doubt that four vertical stripes on the sleeves of clothing serve the source identifying function as strongly. Hence, though we have some doubts about the correctness of the decision on its own terms, we would distinguish *Adidas-America, Inc. v. Payless Shoesource, Inc.*, 546 F. Supp. 2d 1029 (D. Or. 2008), in which the court found the defendant’s use of four stripes on shoes to violate Adidas’s rights. In our view, the stripes on the arms of the clothing depicted above are much more likely to be seen as decorative.

196. Complaint & Demand for Jury Trial, *Lagunitas Brewing Co. v. Sierra Nevada Brewing Co.*, No. 3:15-cv-00153-EDL (N.D. Cal. Jan. 12, 2015). The complaint has since been withdrawn.

197. *Id.* ¶ 11.

variety of different brewers.¹⁹⁸ The term IPA itself is therefore likely generic and unprotectable.

Figure 5.
General Comparison of IPA Labels



Nevertheless, Lagunitas insisted that the “unique ‘IPA’ lettering used in the Lagunitas ‘IPA’ Family of Trademarks has a distinctive serif font, distinctive kerning (or letter spacing), between the ‘P’ and the ‘A’, slightly aged or weathered look, with uneven areas on each of the letters, and the elimination of any periods between the letters.”¹⁹⁹

198. *Id.*; see also *infra* Figure 5.

199. Complaint & Demand for Jury Trial, *supra* note 196, ¶ 16; see also *infra* Figure 6.

Figure 6.
Comparison of IPA Labels (*Lagunitas Brewing v. Sierra Nevada Brewing*)



According to Lagunitas, Sierra Nevada’s new Hop Hunter IPA packaging infringed Lagunitas’ family of IPA marks:

This proposed [Sierra Nevada] design uses all capital, large, bold, black “IPA” lettering in a font selection that is remarkably similar to the iconic Lagunitas design and, indeed, is the central and most prominent feature of the new Sierra Nevada design, emulating the iconic Lagunitas Family of IPA Trademarks. [The design] even uses the kerning between the “P” and the “A” characters that is distinctive to the Lagunitas Family of IPA Trademarks, and in the case of the “neck” label, the proposed Sierra Nevada “IPA” lettering is depicted with the aged or weathered look distinctive to the iconic Lagunitas IPA Family of Trademarks.²⁰⁰

Once we look not just at the plaintiff’s and defendant’s marks, but also at the other uses of IPA, it becomes clear that any protectable element in the IPA mark is extremely narrow. But a jury that considers only infringement might not focus on those limitations unless they are explicitly made part of the strength analysis. Even then the limitations might not exert the force they should, since strength is only one factor in determining likelihood of confusion.

Courts have an especially difficult time managing the scope of complex marks with multiple elements, including trade dress. One

200. Complaint & Demand for Jury Trial, *supra* note 196, ¶ 20.

reason is the greater difficulty defining the mark in these cases. Take, for example, the recent decision in *Reynolds Consumer Products v. Handi-Foil Corp.*²⁰¹ In that case, Reynolds objected to the packaging of a new line of Handi-Foil aluminum foil products.²⁰² Reynolds claimed to own rights in the “overall look, feel and commercial impression of its Reynolds Wrap packaging design.”²⁰³ Specifically, Reynolds identified at least twelve elements that make up this common law trade dress:

(a) the color scheme; (b) the use of the color scheme; (c) the use of prominent lettering within the blue section; (d) the use of silver bands to separate the blue and pink sections; (e) the placement of the quantity information; (f) the inclusion of the “made in the U.S.A.” lettering; (g) the repetition of the color scheme and type on the ends of the boxes; (h) the “press here to lock” feature at the ends of the boxes; (i) the use of graphics with how-to information; (j) the “Lift” graphic; (k) the positioning of the brand name on the top flap of the box; and (l) the color yellow, placement, and text used to caution the consumer.²⁰⁴

Though it noted the maxim that trade dress with more elements tends to have narrower scope, the court never addressed validity or attempted to identify the features that made the Reynolds trade dress protectable.²⁰⁵ It simply engaged in a side-by-side comparison of Handi-Foil’s and Reynolds’s packages, which was enough for the court to conclude that “the similarity between the overall impressions” was “striking.”²⁰⁶

201. No. 1:13-CV-214, 2014 WL 3615853 (E.D. Va. July 18, 2014).

202. *See id.* at *1.

203. *Id.* at *7.

204. *Id.*

205. *See id.* at *8.

206. *Id.* at *9.

Figure 7.
Comparison of Reynolds Wrap and Handi-Foil Packaging Design
(*Reynolds v. Handi-Foil*)



In fact, the court thought the similarities were striking “even if the boxes’ color schemes [were] put aside.”²⁰⁷ What were the damning similarities?

[T]he two boxes both say “non-stick” and “heavy duty.” The Reynolds box says “foods lift right off!” and the Handi-Foil box says “Food Easily Lifts Off!” Both boxes place the (identical) square footage on the right end of the box, with the metric conversions typed neatly below. Additionally, the two-dimensional images above cannot capture the similarity of the boxes’ side and end panels, all of which bear striking semblance. Add to these characteristics the “Made in USA” language on the right portion of the boxes in Example No. 1.²⁰⁸

These elements of similarity, and the court’s belief that Handi-Foil attempted to mimic the Reynolds Wrap box in order to enter the aluminum foil market, convinced the court that the jury’s trade dress verdict was reasonable.²⁰⁹ In fact, according to the court, “any other finding may well have been unreasonable.”²¹⁰

The problem with this conclusion is that the elements of similarity called out by the court were not even elements of the trade dress articulated by Reynolds, very likely because those elements almost

207. *Id.*

208. *Id.* at *8.

209. *See id.*

210. *Id.*

certainly were not things Reynolds was entitled to own. Phrases like “Food Easily Lifts Off” and “Made in the USA” are clearly descriptive, so any rights Reynolds owned in relation to the phrases must have been related to the particular stylization, which Handi-Foil did not copy. When one limits the trade dress in the Reynolds Wrap packaging to only the sorts of things Reynolds is legally entitled to own, the similarities disappear. But because the court treated infringement separately from the question of validity, it reached a clearly incorrect result. This kind of gap is the result of separating validity from infringement and particularly the lack of any integrated mechanism by which to delineate the proper scope of Reynolds’s rights. Indeed, whether any of the similarities ought to have been relevant to a confusion analysis depends on the scope of Reynolds’s rights, an issue which the court never even engaged.

Many courts have responded to this challenge in trade dress cases by requiring claimants to identify the elements of the claimed trade dress.²¹¹ As these courts recognize, it is impossible to determine the scope of a party’s rights without first definitively ascertaining the claimed property. But as the *Reynolds* case demonstrates, articulating the elements of the trade dress is not sufficient because courts are reluctant to identify specifically which of those elements triggers protectability.²¹² Especially with trade dress, courts are often concerned that, by highlighting particular elements, a factfinder might lose track of the totality of the trade dress and “miss an overall similarity.”²¹³

The risk in refusing to highlight particular features is that similarity judgments will be driven by similarity of the unprotectable

211. See, e.g., *Gen. Motors Corp. v. Lanard Toys, Inc.*, 468 F.3d 405, 415 (6th Cir. 2006) (“In requiring a list of ‘discrete elements,’ we are looking to avoid ‘vague and indeterminate references to the “overall appearance” or “look” of plaintiff’s packaging.’” (quoting 1 MCCARTHY, *supra* note 23, § 8:3)); *Yurman Design, Inc. v. PAJ, Inc.*, 262 F.3d 101, 116-17 (2d Cir. 2001) (“[W]e hold that a plaintiff seeking to protect its trade dress in a line of products must articulate the design elements that compose the trade dress.”).

212. See *Kendall-Jackson Winery, Ltd. v. E. & J. Gallo Winery*, 150 F.3d 1042, 1050 (9th Cir. 1998) (“[T]he proper inquiry is not whether individual features of a product are functional or nondistinctive but whether the whole collection of features taken together are functional or nondistinctive.”).

213. *August Storck K.G. v. Nabisco, Inc.*, 59 F.3d 616, 620 (7th Cir. 1995). This is simply an application of the familiar antidissection rule, under which “a composite mark is tested for its validity and distinctiveness by looking at it as a whole, rather than dissecting it into its component parts.” 2 MCCARTHY, *supra* note 23, § 11:27.

features. In theory, a court could guard against this improper influence by requiring a well-constructed survey that used the closest noninfringing substitute as a control, thereby differentiating confusion that resulted from copying of protectable elements from any other form of confusion. But in practice there are several problems with that approach. First, Barton Beebe's empirical research suggests courts don't often rely on surveys, emphasizing them only to support conclusions they have reached for other reasons.²¹⁴ Second, and probably more importantly, one would have to identify the protectable elements of the trade dress and differentiate them from unprotectable elements even to construct the kind of survey we have in mind, in which the closest noninfringing substitute is used as the control.²¹⁵ Courts' reluctance to define the scope of the trade dress therefore infects their ability to evaluate the quality of a control.

Take, for example, *Gucci America, Inc. v. Guess?, Inc.*, in which Gucci claimed that a variety of Guess? products infringed its Repeating GG Pattern trade dress, which consisted of "a pair of inward facing, inverted G's set at the corners of a repeating, diamond-shaped pattern connected by two dots forming straight diagonal lines."²¹⁶ The "Diamond Motif Trade Dress [was] this pattern executed on canvas in a brown/beige colorway, with pinpoint 'shading' within the Gs."²¹⁷

Clearly Guess?'s design shared some of those features, particularly the use of a diamond-shaped pattern connected by Gs, canvas execution in a brown/beige colorway, with pinpoint shading within the Gs.²¹⁸

214. See Beebe, *supra* note 183, at 1641-42.

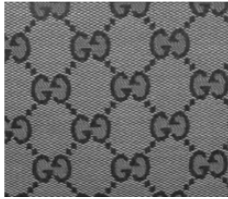
215. See Richard Craswell, "Compared to What?" *The Use of Control Ads in Deceptive Advertising Litigation*, 65 ANTITRUST L.J. 757, 758 (2007).

216. 868 F. Supp. 2d 207, 218 (S.D.N.Y. 2012). Full disclosure: one of us (McKenna) worked as an expert for Guess? in this case. His testimony did not relate to the consumer surveys in the case but instead focused on a marketing report submitted by another expert, which offered a theory of harm to Gucci from Guess?'s sales.

217. *Id.* at 219.

218. See *id.* at 247.

Figure 8.
Repeating GG Pattern (*Gucci v. Guess?*)



*Gucci's Repeating GG
Pattern / Diamond Motif Trade Dress*



Guess's Quattro G Pattern

Thus, standing alone, the two designs no doubt look quite similar—similar enough that one might be concerned about confusion. But comparison of only these two designs would be highly misleading; one or more of those features could be found in many other bags that were on the market at the same time as Guess's. In particular, other bags featured diamond-shaped patterns, canvas execution, and/or the brown/beige colorway:

Figure 9.
Comparison of Bags Sharing Similar Features to Guess's Design
(*Gucci v. Guess?*)



Coach 2006



Christian Dior 2006



*Dooney & Bourke
(since at least 2005)*

In spite of these coexisting designs, the court found “that there [was] a likelihood of confusion between the Diamond Motif Trade Dress and the Quattro G Pattern, when it [was] rendered in brown/beige colorways.”²¹⁹ That conclusion surely was influenced by the court’s finding that Guess? had intentionally copied the Repeating GG Pattern and the Diamond Motif Trade Dress, a finding that was based on the fact that Guess? had sent its manufacturers copies of the Gucci design and asked them to make the Guess? design similar.²²⁰ But the court also was influenced by a survey that perfectly illustrates our concern.²²¹ As shown in Figure 10, Gucci’s survey expert used as the test bag a version of an actual Guess? bag that was modified to remove the central striped element, which Gucci separately claimed as a trademark.²²² The control, however, changed virtually everything—the color, the orientation of the logos on the bag, and even the design of the interlocking four G’s on Guess?’s bag (which, incidentally, is the Guess? Quattro G trademark).²²³

Figure 10.
Gucci Expert Survey (*Gucci v. Guess?*)



Actual Guess? Bag

Mantis Test Bag

Mantis Control Bag

As it turned out, more people believed Gucci was the source of the test bag than the control bag—a result that should surprise no one, given how different the control bag looks.²²⁴ But there is no reason

219. *Id.* at 248.

220. *See id.* at 226-27.

221. *See id.* at 232-33.

222. *See id.*

223. *See id.*

224. Remarkably, the court found infringement despite its conclusion that, combining

to think any confusion regarding the source of the test bag was a result of the protectable aspects of Gucci's design. Indeed, it's entirely possible the respondents simply believed that Gucci is the source of any bag with a brown/beige colorway and logos arranged diagonally.

Perhaps because they are cognizant of the difficulties of managing scope at the infringement stage, courts' concerns about scope sometimes affect their assessments of validity. In some trade dress cases, courts have ratcheted up the secondary meaning requirement, demanding more than the sort of commercial success evidence they routinely accept in word mark cases. In *Yankee Candle*, for example, the court refused to recognize any rights in the common features of Yankee's candle labels, which it described as "essentially a combination of functional and common features," and subjected them to a rigorous secondary meaning analysis.²²⁵ This kind of concern even works at the wholesale level, especially with regard to trade dress. The Supreme Court's *Wal-Mart Stores, Inc. v. Samara Brothers, Inc.* decision, under which product configuration can never be considered inherently distinctive, was based in part on the Court's concerns about overclaiming under a rule that allowed claims to proceed on the basis of alleged inherent distinctiveness.²²⁶

In other cases, courts' concerns about inability to appropriately cabin a party's rights lead them to declare the claimed trade dress entirely invalid because the plaintiff seems to be overclaiming. In *Wallace International Silversmiths, Inc. v. Godinger Silver Art Co.*, for example, the court refused to recognize rights in Wallace's claimed Baroque-style silverware design.²²⁷ It did so on the ground that the design was aesthetically functional, but it seems clear to us that the court believed the real problem was one of scope. According to the court, "if Wallace [had been] able to show secondary meaning in

several surveys, the "maximum level of confusion amongst casual observers in the post-sale setting" was just 5.8 percent. *Id.* at 232.

225. *Yankee Candle Co. v. Bridgewater Candle Co.*, 259 F.3d 25, 42 (1st Cir. 2001); *see also id.* at 44 (requiring more than "evidence of the pervasiveness of the trade dress," noting that anything less would "provide trade dress protection for any successful product").

226. 529 U.S. 205, 213-14 (2000) (noting the lack of clarity in tests of inherent distinctiveness and the "unlikelihood of inherently source-identifying design" and therefore declaring inherent distinctiveness for product design "not worth the candle").

227. 916 F.2d 76, 82 (2d Cir. 1990), *abrogated on other grounds by* *Knitwaves, Inc. v. Lollytogs Ltd.*, 71 F.3d 996 (2d Cir. 1995).

a precise expression of baroque style, competitors might [have been] excluded from using an identical or virtually identical design.”²²⁸ In other words, at the level of “Baroque-style silverware,” Wallace’s design did not have secondary meaning, and allowing Wallace to claim rights at that level would have enabled it to capture designs desired by consumers primarily for their aesthetic value rather than their source designation.²²⁹ If Wallace had claimed narrowly, both problems might have been averted. But since Wallace either would not or could not do so, it had no rights at all.

While, as these cases suggest, courts in trademark cases are sometimes unwilling to recognize any trademark rights when they are concerned about scope, in other cases courts are more comfortable finding noninfringement. That is particularly so when the tool for finding invalidity strikes courts as cutting too broadly, since they understand the relevant invalidity doctrine (most often, functionality) to render the claimed mark unenforceable in any case.

Particularly in aesthetic functionality cases, courts have sometimes been reluctant to invalidate the mark altogether. In *Louboutin v. Yves Saint Laurent*, for example, the Second Circuit seemed to recognize that consumers might want, and competitors might therefore need, red-soled shoes for non-source-related reasons.²³⁰ But rather than finding the red soles unprotectable on that basis, the court fumbled around, suggesting that aesthetic functionality could be considered only *after* resolving all other issues of protectability and confusion,²³¹ and only through a “fact-specific” inquiry and never by “a *per se* rule.”²³² The court then dodged the question entirely by construing Louboutin’s mark narrowly to cover only contrasting red outsoles and not all-red shoes like those Yves Saint Laurent sold.²³³ Rather than invalidating the mark, it limited the scope of Louboutin’s rights and found noninfringement.²³⁴

But unlike *Louboutin*, in which the court found another route to deny recovery to the plaintiff, in many cases courts’ concerns about

228. *Id.*

229. *Id.*

230. *See* Christian Louboutin S.A. v. Yves Saint Laurent Am. Holding, Inc., 696 F.3d 206, 227-28 (2d Cir. 2012).

231. *See id.* at 216-17.

232. *Id.* at 222-23.

233. *See id.* at 228.

234. *See id.*

the consequences of using validity doctrines leave them without a tool even to avoid enforcement against a particular defendant who has not done anything the law legitimately prohibits. In *Automotive Gold, Inc. v. Volkswagen of America, Inc.*, for example, the defendant wanted to use the Volkswagen (VW) logo on key chains and license plate covers, which it plausibly contended consumers wanted to match their cars.²³⁵ Though the court acknowledged that this was possible, it could not bring itself to find that the VW logos were functional, largely because it felt that finding would compel the conclusion that VW owned no rights in the VW logos for any purposes, including when used to designate the source of Volkswagen's vehicles.²³⁶ Thus, the court found the defendant liable for a use that trademark law should not prohibit because the tool it had available swept too broadly.²³⁷

Similarly, in *Rosetta Stone Ltd. v. Google, Inc.*, the Fourth Circuit held that Google could face trademark liability for the automatic operation of its AdWords feature, which showed people ads they might be interested in based on the search terms they typed into Google.²³⁸ The district court held that while Google technically "used" the mark, its use was a functional one—for the purpose of returning an automated result.²³⁹ But the Fourth Circuit rejected that conclusion, deciding that the tool the district court used to achieve that end—functionality—was an all-or-nothing validity rule and could not be

235. See 457 F.3d 1062, 1065 (9th Cir. 2006).

236. See *id.* at 1067 ("Consumers sometimes buy products bearing marks such as the Nike Swoosh, the Playboy bunny ears, the Mercedes tri-point star, the Ferrari stallion, and countless sports franchise logos, for the appeal of the mark itself, without regard to whether it signifies the origin or sponsorship of the product.").

237. One of us has previously argued that aesthetic functionality need not have this all-or-nothing character—that courts could find that particular uses were for their aesthetic function and thereby insulate the defendant from liability without declaring the mark entirely invalid. See McGeeveran & McKenna, *supra* note 53; McKenna, *supra* note 97; cf. Justin Hughes, *Cognitive and Aesthetic Functionality in Trademark Law*, 36 CARDOZO L. REV. 1227, 1230 (2015) (arguing that aesthetic functionality is better understood as an effort to deny trademark owners control over features that provoke a cognitive or psychological response in consumers). Lemley has argued that such uses (for nontrademark purposes) should be deemed noninfringing because only trademark uses can infringe. See *generally* Dogan & Lemley, *supra* note 94. Although these arguments differ in some important respects, both are efforts to delineate the scope of trademark rights for marks that pass the validity threshold.

238. 676 F.3d 144, 156 (4th Cir. 2012).

239. See *id.* at 153; cf. Hughes, *supra* note 237, at 1230.

used to calibrate the scope of Rosetta Stone's rights in its mark.²⁴⁰ Similarly, in *Fleischer Studios*, the Ninth Circuit originally held that a defendant could not be liable for putting a picture of Betty Boop on shirts because it was merely using the image of Betty Boop and not the trademark as such, and trademark law is not supposed to give ownership over the image itself.²⁴¹ The court subsequently reversed itself, however, again concluding that the tool on which it had previously relied—the functionality doctrine—could not be used to calibrate the proper scope of the plaintiff's right.²⁴²

We have argued elsewhere that trademark law has expanded inappropriately in various respects and is in need of limiting principles.²⁴³ One of the reasons courts find liability in cases that extend well beyond the proper scope of trademark rights is that the limiting doctrines that are available turn out to be all-or-nothing validity doctrines. Courts are (properly) reluctant to invalidate a plaintiff's mark altogether, and they are unwilling or unable to limit the scope of that mark in cases where it has been asserted against a defendant that is not doing something the law actually forbids.

4. Design Patent

Unlike utility patents, design patents are supposed to protect only the ornamental features of a design, not the functional aspects that make the device work better.²⁴⁴ If a design patent owner tries to assert control over a functional aspect of the design, she is

240. See *Rosetta Stone*, 676 F.3d at 162 (“Once it is determined that the product feature—the word mark ROSETTA STONE in this case—is not functional, then the functionality doctrine has no application, and it is irrelevant whether Google’s computer program functions.”).

241. See *Fleischer Studios, Inc. v. A.V.E.L.A., Inc.*, 636 F.3d 1115, 1124 (9th Cir. 2011), *rev’d on reh’g*, 654 F.3d 958 (9th Cir. 2011).

242. See *Fleischer Studios, Inc. v. A.V.E.L.A., Inc.*, 654 F.3d 958, 960 (9th Cir. 2011). For a general discussion on the Betty Boop case, see Irene Calboli, *Betty Boop and the Return of Aesthetic Functionality: A Bitter Medicine Against “Mutant Copyrights”?*, 36 EUR. INTEL. PROP. REV. 80 (2014).

243. See generally Lemley & McKenna, *supra* note 67.

244. See *PHG Techs., LLC v. St. John Cos.*, 469 F.3d 1361, 1366 (Fed. Cir. 2006). Unlike trademark law, which is designed to protect identifications of source, design patent is intended to protect aesthetics directly. So, there is no aesthetic functionality doctrine in design patent as there is in trademark. See, e.g., J.H. Reichman, *Design Protection and the New Technologies: The United States Experience in a Transnational Perspective*, 19 U. BALT. L. REV. 6, 8-10 (1989); Michael Risch, *Functionality and Graphical User Interface Design Patents*, 17 STAN. TECH. L. REV. 53, 55 (2013).

improperly seeking to expand the scope of her right beyond what the law is supposed to give.

One scope problem with design patents parallels the one just discussed in trademarks: functionality. The functionality doctrine is intended to stop this sort of overreaching. Unfortunately, it is not up to the task. Because functionality is a validity doctrine, not an infringement doctrine or a defense, a finding that a design patent is functional means no protection at all for the design, even for its ornamental features.²⁴⁵ Courts, unwilling to invalidate a design altogether merely because part of the designed device is functional, have responded by narrowing the design patent functionality doctrine almost to the point of nonexistence.²⁴⁶ In design patent law, a design is functional—and thus unprotectable—only if there is no other possible way to achieve the same result.²⁴⁷ And a functional design is unprotectable only if it is “governed solely by function.”²⁴⁸ Needless to say, that almost never happens.²⁴⁹ But if the design is treated as nonfunctional for validity purposes, the fact that it does function is then ignored when it comes to infringement and defenses, because functionality is a validity doctrine, not an infringement doctrine. The result is that it is all too possible that a design patent owner can capture scope she was never meant to have, preventing her competitors from copying functional rather than ornamental features of a product.²⁵⁰

245. *See Power Controls Corp. v. Hybrinetics, Inc.*, 806 F.2d 234, 238 (Fed. Cir. 1986).

246. *See Mark P. McKenna & Katherine J. Strandburg, Progress and Competition in Design*, 17 STAN. TECH. L. REV. 1, 2 (2013); Risch, *supra* note 244, at 54.

247. *See, e.g., Seiko Epson Corp. v. Nu-Kote Int'l, Inc.*, 190 F.3d 1360, 1368 (Fed. Cir. 1999) (holding that a design is functional only if it is “the only possible form of the article that could perform its function”); *L.A. Gear, Inc. v. Thom McAn Shoe Co.*, 988 F.2d 1117, 1123 (Fed. Cir. 1993); *Avia Grp. Int'l, Inc. v. L.A. Gear Cal., Inc.*, 853 F.2d 1557, 1563 (Fed. Cir. 1988). This is a much narrower standard than functionality in trademark law, which prevents protection if something affects the cost or quality of the article being sold. *See Traffix Devices, Inc. v. Mktg. Displays, Inc.*, 532 U.S. 23, 32 (2001).

248. *Seiko Epson*, 190 F.3d at 1368; *see In re Carletti*, 328 F.2d 1020, 1022 (C.C.P.A. 1964).

249. For a discussion of the few cases invalidating design patents on functionality grounds in recent years, see Risch, *supra* note 244, at 92.

250. *See Egyptian Goddess, Inc. v. Swisa, Inc.*, 543 F.3d 665, 672, 682-83 (Fed. Cir. 2008) (en banc). The prior test had permitted a defendant to avoid infringement if he could show that he used only the unprotectable elements of the design. *See Lee v. Dayton-Hudson Corp.*, 838 F.2d 1186, 1188 (Fed. Cir. 1988). The “ordinary observer” test does not seem to contemplate such a “practicing the functional art” defense.

We could solve that problem at the infringement stage if we applied functionality—or something like it—as a legal scope filter to ensure that defendants were not held liable based on similarities of functional rather than ornamental features. But in fact the functionality problem is compounded by a second gap in scope—the test for infringement. The basic test for infringement of a design patent is the “ordinary observer” test, which assesses “whether an ordinary observer, familiar with the prior art, would be deceived into thinking that the accused design was the same as the patented design.”²⁵¹ But that test has long been supplemented by a requirement that the defendant have appropriated the thing that makes the design new and different from the prior art.²⁵² If the defendant’s use didn’t include the novel feature(s) in the patentee’s design, it couldn’t infringe.²⁵³

In 2008, in *Egyptian Goddess v. Swisa*, the Federal Circuit changed the law of design patents to eliminate the point-of-novelty test for infringement.²⁵⁴ Instead, the court said, the focus must be on the overall appearance of the design as a whole, not on “small differences” at the point of novelty.²⁵⁵ Novelty still matters, but now only as a defense that must generally be assessed separately from infringement.²⁵⁶ And subsequent Federal Circuit cases have used *Egyptian Goddess* as precedent in concluding that point of novelty is no longer the test for the invalidity of a design patent either.²⁵⁷

251. *Egyptian Goddess*, 543 F.3d at 672.

252. *See id.* at 670-71.

253. Portions of these three paragraphs are adapted from Lemley, *supra* note 24, at 1270-71.

254. 543 F.3d at 678.

255. *Id.* at 677.

256. *See id.* at 678. The court did, however, take some steps to try to sneak back in the prior art comparison it had just rejected, declaring that the ordinary observer was henceforth one who was “familiar with the prior art.” *Id.* at 681. But the separation of the infringement and validity inquiries creates an even bigger problem here than it does in utility patent law. To infringe a utility patent, the defendant must include every element of the patented claim, and so must necessarily include the thing that makes the invention patentable. After *Egyptian Goddess*, by contrast, design patent law no longer has such a requirement. A design patent can now be infringed even by a product that lacks the new feature encompassed by the patent as long as an ordinary observer would think the two were substantially the same overall.

257. *See, e.g., Int’l Seaway Trading Corp. v. Walgreens Corp.*, 589 F.3d 1233, 1237-38 (Fed. Cir. 2009). For a discussion of this extension, see Marta Kowalczyk, *Design Patent Infringement: Post-Egyptian Goddess*, 2010 U. ILL. J.L. TECH. & POL’Y 239, 250-56.

As long as an ordinary observer would find the design of the accused product substantially the same as the patented design, the fact that such a conclusion is driven by similarities in features that already exist in the prior art doesn't defeat a finding of infringement. It might or might not create a defense that the patent is invalid for anticipation, though again that seems to depend on what an ordinary observer would think when comparing the patented design and the prior art.²⁵⁸ And in any event that defense is evaluated in isolation from infringement, not as part of an integrated analysis of the proper scope of the design patent.

258. See *Int'l Seaway*, 589 F.3d at 1238. The Federal Circuit backed off that "ordinary observer" test in favor of obviousness, which seems inconsistent with *KSR International Co. v. Teleflex, Inc.*, 550 U.S. 398, 420 (2007), and *High Point Design LLC v. Buyers Direct, Inc.*, 730 F.3d 1301, 1313 (Fed. Cir. 2013) (holding that the standard for obviousness was the ordinary designer, not the ordinary observer). Because the ordinary designer is presumably more knowledgeable about design than the ordinary observer, the result of this change may be to render more design patents invalid, though the obviousness standard is problematic in other respects.

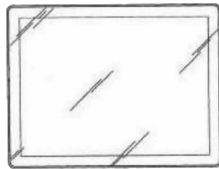
Figure 11.
Comparison of the Prior Art, Apple iPad, and Samsung Tablets
(*Apple v. Samsung*)



Prior Art 1



Prior Art 2



Apple iPad



Samsung Tablet

To see the difference, consider the *Apple, Inc. v. Samsung Electronics Co.* litigation.²⁵⁹ Apple's design patent for the iPad is on the bottom left; Samsung's accused tablet is on the bottom right. An infringement analysis conducted in isolation asks whether an ordinary observer would consider the Samsung design too close to the patented one.²⁶⁰ In fact, however, Apple is not the first to come up with a computer tablet that is square with rounded corners.²⁶¹ The top two images in Figure 11 are in the prior art. A comparison that includes the prior art would emphasize the narrowness of what Apple contributed to the design and ask whether Samsung appropriated whatever Apple's patent added to the prior art. The move away from that integrated scope question to the ordinary observer question creates a gap—a risk that Apple will prevail because the jury

259. See 678 F.3d 1314, 1319 (Fed. Cir. 2012).

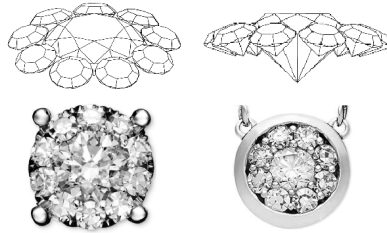
260. See *id.* at 1320.

261. See *id.* at 1329-30.

finds similarity in the parts of the Apple design that are not protectable: the things Apple took from the prior art. At the very least, it puts a great deal of reliance on the jury to use their presumed “familiarity with the prior art” to implicitly exclude from consideration of similarity features that do not owe their origin to Apple.

Similarly, in *Prestige Jewelry International, Inc. v. BK Jewellery HK*, BK’s patented design featured a “relatively large full-cut center diamond” that was tightly surrounded by “nine peripheral single-cut diamonds.”²⁶² BK alleged that Prestige infringed its design by selling the Unity cluster-top jewelry products, one version of which “featured a full-cut center diamond surrounded by eight or ten relatively smaller single-cut diamonds.”²⁶³

Figure 12.
Unity Images (*Prestige Jewelry v. BK Jewellery*)

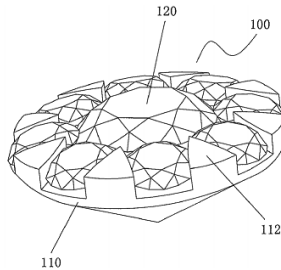


BK was, of course, not the first one to make jewelry with a larger diamond surrounded by smaller diamonds. One piece of prior art (“the Lin Patent”), for example, disclosed a larger center diamond surrounded by eight smaller, full-cut diamonds.

262. No. 1:11-cv-02930-LAP-DCF, at *1 (S.D.N.Y. Oct. 15, 2014) (Sept. 15, 2014 order denying summary judgment). “Full-cut diamonds have more facets than single-cut diamonds and typically produce a greater sparkling effect.” *Id.* at *5.

263. *Id.* at *10. Prestige also sold a version of the Unity jewelry that included only full-cut diamonds.

Figure 13.
The Lin Patent No. 7,762,104 (*Prestige Jewelry v. BK Jewelry*)



Another piece of prior art, the “Edwardian cluster ring,” “included nine peripheral stones and was practiced at least as early as the Edwardian era, approximately 1901-1915.”²⁶⁴ A proper assessment of BK’s design would have focused particularly on the use of a relatively larger center diamond surrounded by nine smaller *peripheral single-cut* diamonds. That *Prestige* surrounded a larger central diamond with eight or ten (or even nine) smaller peripheral diamonds ought not be enough to say it is infringing. But the court expressly refused to identify the point of novelty, assuming the jury could see for itself the novel features of the design and know not to credit similarities in anything but those novel features.²⁶⁵ We seriously doubt that.

A final potential scope gap in design patents involves the assertion of design patents against uses of the design on different products than the plaintiff expected. The design patent claim takes the form of an image. That image is normally representative of a product, but increasingly it is an image or an icon or brand that could be applied to a product.²⁶⁶ When the defendant uses the image in a completely different context (say, reproducing the design in a book or using it on a bumper sticker), the question is properly whether

264. *Id.* at *14.

265. *See id.* at *30-31.

266. *See, e.g.,* Jason J. Du Mont & Mark D. Janis, *Virtual Designs*, 17 STAN. TECH. L. REV. 107, 109 (2013). For an argument that design patents should not be extended beyond the shape of a specific type of product to images that can be applied to other objects, see generally Sarah Burstein, *The Patented Design*, 83 TENN. L. REV. 161 (2015).

the defendant's use is within the lawful scope of the design patent. But is that a question of validity or one of infringement? Design patent law thinks about the issue as one of validity—is a patent appropriately limited to application to a product? But it is perhaps more plausibly a scope question. The problem with Apple suing someone who puts their logo on a bumper sticker isn't that the logo patent is invalid, but that they are using that patent to control behavior over which we don't want to give them control, just as we don't want to give Volkswagen the right to control the making of keychains. Without a way of calibrating the scope of the design patent, we are left with a choice between invalidating the patent altogether or allowing the IP owner to prevent a use that design patent law was never intended to prevent.²⁶⁷

Allowing enforcement of a design patent against these kinds of uses would also run afoul of the maxim that “[t]hat which infringes, if later, would anticipate, if earlier.”²⁶⁸ In design patent law, “[t]he scope of the prior art is not the universe of abstract design and artistic creativity, but designs of the same article of manufacture or of articles sufficiently similar that a person of ordinary skill would look to such articles for their designs.”²⁶⁹ Under that standard, a bumper sticker is unlikely to count as relevant prior art against a GUI design. Yet if the design patent isn't field-limited, then it's no longer true that that which infringes if later would invalidate if earlier. The scope of the design patent on the infringement side is no longer constrained by what differentiates the invention from the prior art.

III. TOWARDS AN INTEGRATED DOCTRINE OF IP SCOPE

A. *The Need for an Integrated Scope Determination*

What all of the examples in Part II have in common is a disconnect between legal doctrines that allows an IP owner to claim something that the law should not—and indeed purports not to—give

267. For an argument that design patents should not extend to representations of the design or its use on different products than the patentee's, see Burstein, *supra* note 266.

268. *Int'l Seaway Trading Corp. v. Walgreens Corp.*, 589 F.3d 1233, 1239 (Fed. Cir. 2009) (alteration in original) (quoting *Peters v. Active Mfg. Co.*, 129 U.S. 530, 537 (1889)).

269. *Hupp v. Siroflex of Am., Inc.*, 122 F.3d 1456, 1462 (Fed. Cir. 1997).

her. The proper scope of an IP right is not set by some god or dictated by immutable natural law. It is a function of the purposes of the IP regime, what the IP claimant has contributed to the world, and the uses that must remain open in order that the IP regime not unduly interfere with social welfare.²⁷⁰ The point of complex validity rules, infringement rules, and defenses is to get as close as possible to that proper scope.

Because the law has separated validity, infringement, and defenses, generally considering each of them in isolation from the others, IP owners can and regularly do seek to expand the scope of their IP rights beyond permissible bounds. The law sometimes reacts by invalidating the IP right altogether because of the overbreadth of the plaintiff's claim. But courts and especially juries tend to be reluctant to strike down an IP right altogether for overclaiming.²⁷¹ And even though invalidating the right can be the better of the two options available to the court, that result is often not ideal, since invalidating the right means that it cannot be asserted even against infringements that are concededly within the proper, more limited scope of the IP right.

The separation of validity, infringement, and defenses has turned IP litigation into something akin to a game of blackjack. Whatever the hand the IP owner has been dealt, they have an incentive to add to it—increase the scope of the right—as much as possible. But if they go too far—over twenty-one—the entire hand is thrown out. Whatever the merits of blackjack as a card game, it is not a reliable way to get IP regimes to produce the right answer, for it often leads to over- or underinclusive rights.

IP law needs a way to expressly consider scope in a single, integrated proceeding. Courts should coordinate validity, infringement, and defense proceedings in some way so that both the fact of overreaching and its potential consequences become clear to the parties and the court before trial. Our aim should be to find the proper scope for an IP right. And that can't be done if we consider validity, infringement, and defenses in isolation. Courts must be willing to think about these doctrines in an integrated fashion.

270. See Lemley & McKenna, *supra* note 67, at 183-84.

271. See *Rhine v. Casio, Inc.*, 183 F.3d 1342, 1345 (Fed. Cir. 1999) (reasoning that “claims should be construed to sustain their validity” if possible).

That, in turn, means they will have to understand and confront the purposes of the IP doctrines, rather than viewing those doctrines as rules to follow wherever they might lead.

We think many of copyright and trademark law's current features are a consequence of the lack of an integrated scope mechanism. Existing doctrines reflect tradeoffs between validity, infringement, and defenses as mechanisms to manage scope. Patent law chooses a high bar for validity, and *Markman* hearings set a defined meaning that controls at the infringement stage.²⁷² As a result, patent law has less work to do to manage scope at the infringement stage, and it has relatively less need for defenses. Copyright is at the other end of the spectrum, setting a low bar for validity and punting the great majority of scope management to infringement, defenses, and exceptions. Trademark is in between, but closer to copyright than to patent: it has some true validity doctrines, but it still tries to manage scope at the infringement stage much more than patent. This means that copyright and trademark have a tendency to hand many important scope questions to the jury.²⁷³

That allocation would be fine if we thought courts had good tools for managing scope at the later stages of a case. But they don't. In copyright cases, courts do analytic dissection for purposes of determining copying but not for purposes of infringement.²⁷⁴ Trademark law doesn't even have (explicitly at least) anything analogous to analytic dissection; it simply counts on the likelihood of confusion factors to set the scope.²⁷⁵ And although these doctrinal structures can and do generate errors in both directions, we think they more commonly result in overbroad protection for works and trademarks that deserve only very narrow protection, if any.

B. *Markman as a Scope Proceeding*

There is one area of modern IP law that treats scope in a more-or-less integrated way: utility patent law. For the past twenty years, patent cases have included a pretrial proceeding in which the courts

272. See *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 374 (1996).

273. See generally Balganesch et al., *supra* note 155 (discussing the different questions answered by judges and juries during copyright trials).

274. See Lemley, *supra* note 112, at 736-38.

275. See *supra* Part II.C.3.

construe the claims of the patent, defining what the legal definition of the patent is—and therefore what the patent does and does not cover—in the specific context of the dispute between the parties. These so-called *Markman* hearings do what virtually no other IP proceeding does: they define the breadth of the patent right in a way that is thereafter used for both validity and infringement.²⁷⁶ In theory, therefore, they prevent the nose of wax problem that has beset assessments of scope in IP law.

Markman hearings frequently result in summary judgment for one side or the other—once the court has decided the proper scope of the patent, it is rare that there is a factual dispute as to whether the defendant has infringed the patent.²⁷⁷ And while post-*Markman* disputes about validity are more common, many of those can be resolved once the court has decided what the patent claim does and does not cover. Further, the advent of *Markman* hearings led directly to the closing of one scope gap—the temporal difference between patent validity and patent infringement. Before *Markman*, courts assessed patent validity based on how the claim would be understood as of the time of invention, or perhaps the filing date, but assessed infringement based on the understanding of the term as of the date of infringement.²⁷⁸ But once courts started construing claim terms in a single integrated proceeding, they rebelled against the idea that the same claim should have different meanings at different times.²⁷⁹

Markman hearings are not by any means a perfect expression of the scope principle. In particular, there are two problems in practice with *Markman* hearings as scope vehicles. First, although courts are supposed to construe claims to preserve their validity where possible,²⁸⁰ the Federal Circuit has discouraged courts from thinking about the validity of the claims at all when deciding the meaning of

276. Named for *Markman*, 517 U.S. at 370, 384-87, which created them. There are few substantive defenses in patent law apart from invalidity, but when they come into play, they too are governed by the definition of the invention set out in the *Markman* proceeding.

277. See John R. Allison & Mark A. Lemley, *The (Unnoticed) Demise of the Doctrine of Equivalents*, 59 STAN. L. REV. 955, 958 (2007).

278. See Lemley, *supra* note 8, at 105-10.

279. See *id.* at 110-11. Lemley endorses this integrated approach. See *id.* at 112-15.

280. See, e.g., *Rhine v. Casio, Inc.*, 183 F.3d 1342, 1345 (Fed. Cir. 1999) (holding that claims should be so construed, if possible, as to sustain their validity).

the claim terms at issue in the case.²⁸¹ This is an instance of the “different decisionmakers” problem.²⁸² Courts seem unwilling to assess validity pretrial because it is (sometimes) an issue for the jury at trial.²⁸³ The result is that, even at *Markman* hearings, patentees may seek to unduly broaden the scope of the patent and take their chances later on with invalidity. The situation is still better than it would be without the hearing, because the claim is given a single scope for both validity and infringement purposes. But it does not avoid the blackjack problem altogether.

Second, *Markman* hearings focus on the words of the claims written by patent lawyers to try to define the invention rather than on the invention itself.²⁸⁴ As one of us has argued elsewhere, allowing words about inventions to become the focus rather than the inventions themselves has led to a variety of mischief, from making overclaiming easier to introducing ambiguities in the scope of the patent that have nothing to do with what the patentee actually invented, but only with uncertainty in the meaning of the words chosen to describe them.²⁸⁵ And this has also had the perverse effect of dividing—rather than unifying—the analysis of a patent claim, focusing on short phrases and individual words rather than the claim as a whole. The result has arguably been to “fraction” patent law, throwing up a large number of hurdles to the patentee rather than evaluating the scope of the claims as an integrated whole.²⁸⁶

But these failures of the *Markman* process should not overshadow its central promise: getting the court to think about the scope of the patent right, rather than just whether it is valid or infringed. So while we strongly disagree with Larry Solum and TJ Chiang that there is anything like an acceptable amount of certainty in the meaning we assign to words in claim construction,²⁸⁷ we very much

281. See *Phillips v. AWH Corp.*, 415 F.3d 1303, 1327-28 (Fed. Cir. 2005) (en banc) (dismissing that canon as one “of limited utility in any event” and saying that “we have not applied that principle broadly”).

282. See *supra* Part II.B.1.

283. For an argument that validity should not necessarily go to the jury, see generally Lemley, *supra* note 72 (explaining that jury decisions on the validity of patents are a relatively new phenomenon and something unique to the American court system).

284. See *Burk & Lemley, supra* note 10, at 1744.

285. See *id.* at 1762.

286. See Mark A. Lemley, *The Fractioning of Patent Law*, in *INTELLECTUAL PROPERTY AND THE COMMON LAW* 504, 509 (Shyamkrishna Balganesh ed., 2013).

287. See generally Tun-Jen Chiang & Lawrence B. Solum, *The Interpretation-Construction*

agree that the *Markman* hearing is at its base a dispute about the scope of the legal right being granted, not simply about the meaning of words. That's why claim construction failed when courts tried to do it in a vacuum, without the context of the accused device.²⁸⁸

And unlike any other area of IP law, patent law for the past twenty years has set out to establish that scope directly, rather than just applying legal doctrines that depend on an implicit and changing assumption about the scope of the patent. Although *Markman* hearings have focused on the wrong mechanism for measuring scope—the words used by lawyers to describe the invention, rather than the invention itself—they are at least trying to determine the scope of the invention in a single proceeding.

C. Adapting the Scope Proceeding for Other IP Rights

Copyright, trademark, and design patent should seek to emulate patent law in adopting an integrated, pretrial determination of the proper scope of the IP right. This doesn't mean these other areas of IP should adopt claims or claim construction verbatim.²⁸⁹ Copyrights

Distinction in Patent Law, 123 YALE L.J. 530, 534 (2013) (arguing that “uncertainty in claim application most typically arises because judges have core policy disagreements about the underlying goals of claim construction” rather than because judges acknowledge different linguistic meanings). As Zahr Said points out, it may be the very transparency into the scope determination that *Markman* provides that makes the uncertainty of the process apparent. See Said, *supra* note 8, at 519-20. Copyright and trademark scope determinations are no more certain; we may simply not see the uncertainty because we don't think expressly about scope at all.

For a different critique of Chiang and Solum, see generally Camilla Hrdy & Ben V. Picozzi, *Claim Construction or Statutory Construction? A Response to Chiang & Solum*, 124 YALE L.J. F. 208 (2014), <http://www.yalelawjournal.org/forum/claim-construction-or-statutory-construction> [<https://perma.cc/G6YF-D49W>]; Oskar Liivak, *The Unresolved Interpretive Ambiguity of Patent Claims: A Response to Solum and Chiang 1-7* (Feb. 3, 2015) (unpublished manuscript), <http://ssrn.com/abstract=2559958> [<https://perma.cc/DC53-QWF5>].

288. *Compare* *Vitronics Corp. v. Conceptronc, Inc.*, 90 F.3d 1576 (Fed. Cir. 1996) (describing claim construction as a legal question that could be resolved in the abstract), *with* *Wilson Sporting Goods Co. v. Hillerich & Bradsby Co.*, 442 F.3d 1322, 1326-27 (Fed. Cir. 2006) (rejecting the *Vitronics* approach: “[T]he legal function of giving meaning to claim terms always takes place in the context of a specific accused infringing device or process.... [K]nowledge of that product or process provides meaningful context for the first step of the infringement analysis.”). Most courts follow *Wilson Sporting Goods* today. See, e.g., *Kaneka Corp. v. Xiamen Kingdomway Grp. Co.*, 790 F.3d 1298, 1303 (Fed. Cir. 2015).

289. Jeanne Fromer argues that copyright law already has a de facto system of claiming, though it operates centrally rather than peripherally and by example rather than by characteristic. See Fromer, *supra* note 13, at 721. But she resists the idea of building a claiming

and trademarks typically don't have claims at all,²⁹⁰ and design patents have visual claims that aren't necessarily suitable to the sort of verbal clarification that is the heart of the *Markman* hearing.²⁹¹ Utility patent law also has some structural advantages that make it especially well suited to the kind of integrated scope proceeding we described. By comparison to copyright and trademark, and even by comparison to design patent, utility patent law has a relatively high bar for validity, and courts therefore don't have to work as hard to manage scope at the infringement. Certainly there are important questions in patent cases about the meaning of the claims, and as we argued above, that scope determination can't be disentangled from infringement. But patent law has the significant advantage of the "all elements" rule, under which an accused invention infringes only if it contains every element of the claimed invention.²⁹² That rule necessarily draws a factfinder's attention to each and every limitation of the claim in a way that copyright and trademark law's infringement rules do not. And because the patent infringement standard requires identity between the accused device and the claimed elements of the patented invention, it's relatively easy to determine infringement once we know the meaning of the claims.²⁹³ Both copyright and trademark law are more expansive here in that the defendant's work or mark can be more dissimilar and still be infringing. As a result, copyright and trademark still

procedure more explicitly into copyright law. *See id.* at 793-94.

290. We say "typically" because several circuits now require a trade dress claimant to identify the particular elements of the claimed trade dress. *See, e.g.,* Gen. Motors Corp. v. Lanard Toys, Inc., 468 F.3d 405, 415 (6th Cir. 2006) ("In requiring a list of discrete elements, we are looking to avoid vague and indeterminate references to the overall appearance or look of plaintiff's packaging."); Yurman Design, Inc. v. PAJ, Inc., 262 F.3d 101, 116-17 (2d Cir. 2001) ("[W]e hold that a plaintiff seeking to protect its trade dress in a line of products must articulate the design elements that compose the trade dress.>").

291. *See* Egyptian Goddess, Inc. v. Swisa, Inc. 543 F.3d 665, 679-80 (Fed. Cir. 2008) (rejecting the need for a verbal claim construction of a visual claim in a design patent case). As the Supreme Court has explained, a design patent claim "is better represented by [an] illustration than it could be by any description, and a description would probably not be intelligible without the illustration." *Dobson v. Dornan*, 118 U.S. 10, 14 (1886).

292. The doctrine of equivalents obviously loosens this requirement a bit, but the Supreme Court has constrained that doctrine too, requiring equivalency with respect to each claim element. *See* Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 29 (1997). And as we note below, the doctrine of equivalents also builds on another scope doctrine by prohibiting parties from claiming equivalents that cover the prior art.

293. *See id.* at 35.

have more work to do after determining the protectable aspects of the plaintiff's property.

Nevertheless, we think courts should have an opportunity to assess the proper legal scope of copyrights, trademarks, and design patents, just as they do patents. Defining the scope of the IP right is critical for assessing all aspects of the case, and it is something that is best done by the judge before the substantive merits of the case are resolved. The most straightforward way to assess that scope is by identifying in a clear way what aspects of the IP right are protectable. Most importantly, the scope determination must be done in an integrated proceeding that considers what the plaintiff is entitled to claim as well as the differences between the plaintiff's and defendant's products and any available defenses. And that proceeding must happen before a judge, not a jury, except in rare cases. Only a judge is equipped to evaluate the full range of legal doctrines that determine the scope of an IP right.²⁹⁴

That doesn't mean we would take all infringement or validity issues away from the jury, even if we could. Rather, the point is that before courts give one of those issues to the jury—or even resolve it in isolation on a summary judgment motion—they need to engage in a reality check to make sure that the thing that might allow a factfinder to find infringement is something the law protects. Such a determination shouldn't be delegated to a jury unversed in the tradeoffs inherent to the IP system.

There are examples outside the *Markman* context of courts engaging in this kind of integrated scope analysis. The doctrine of equivalents in utility patent law has an integrated scope doctrine of

294. In an important article, Zahr Said argues for an interpretive regime in copyright akin to patent's *Markman* hearing. See Said, *supra* note 8, at 473-74. She affirmatively counsels that courts take a formalist approach to interpreting the scope of copyrights, arguing that doing so will "offer[] defendants a more predictable and streamlined way to cut off litigation pre-trial." *Id.* at 519. We are more skeptical than Said about the virtues or workability of formalism in a fact-bound area like copyright. But we do think she is on the right track to suggest a pretrial interpretive hearing before a judge in which the court has the power to determine the scope of the copyright in an integrated fashion. While it is theoretically true that integrating the entire proceeding before the jury could have the same effect, we don't think juries are sufficiently versed in the legal rules that they could make a scope determination properly. They would also not make it transparently, which would make it difficult for an appellate court to review the scope assessment. And in any event, the rules of copyright and trademark currently preclude giving the jury the sorts of information they would need to make an informed scope decision.

a sort in the “prior art” defense. While, as we have seen, practicing the prior art is no defense to a claim of *literal* patent infringement, it is a defense to infringement under the doctrine of equivalents.²⁹⁵ The doctrine of equivalents creates a sort of penumbra around the scope of the patent claim, allowing patentees to sue people who make products that are only insubstantially different from the patented invention.²⁹⁶ But the doctrine of equivalents does not permit a patentee to use the doctrine to expand the scope of her claims to such an extent that it would cover things in the prior art. To avoid that problem, patent courts will construct a “hypothetical claim” that would cover the defendant’s product, and ask whether that hypothetical claim would also cover items in the prior art.²⁹⁷ If it would, the patentee is overclaiming and the effort to apply the doctrine of equivalents is rejected.²⁹⁸ The prior art defense in patent’s doctrine of equivalents offers an example of an explicit scope decision that is nonetheless something short of a full-blown *Markman* hearing.

While we don’t suggest written claims for copyrights, trademarks, or design patents (and we even worry about how well they work with utility patents), we think courts will often have to express in words exactly what is different and protectable about the IP right.²⁹⁹ Doing so will often occur in connection with a motion for summary judgment, though we would be more comfortable with a separate proceeding, as in patent law, because validity and infringement often won’t both come before the court on summary judgment. But

295. See *Wilson Sporting Goods Co. v. David Geoffrey & Assocs.*, 904 F.2d 677, 683-85 (Fed. Cir. 1990), *overruled on other grounds by* *Cardinal Chem. Co. v. Morton Int’l, Inc.*, 508 U.S. 83 (1993).

296. On the doctrine of equivalents in patent law, see *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 535 U.S. 722, 731-34 (2002); *Warner-Jenkinson*, 520 U.S. at 21; *Graver Tank & Mfg. Co. v. Linde Air Prods. Co.*, 339 U.S. 605, 607-09 (1950).

297. See *Wilson Sporting Goods*, 904 F.2d at 684.

298. See *id.* at 685.

299. Some have expressed to us the concern that dissecting and describing the works can miss the heart of the similarities between them. That is clearly the motivation behind the “total look and feel” approach. But we worry that the law errs too much in the opposite direction, protecting an amorphous look and feel, see *Roth Greeting Cards v. United Card Co.*, 429 F.2d 1106, 1110 (9th Cir. 1970), even when the things that make a work seem similar are not the things IP law intends to protect. IP owners should be able to demonstrate a synergistic combination of known elements in appropriate cases, but that cannot be an excuse for simply handing the two works to the jury with no guidance as to what can and cannot constitute infringement.

even if a jury must resolve a factual issue, courts should be communicating clearly to juries which similarities can—or, as might be easier in many cases, *cannot*—be the basis for a finding of infringement.

Sometimes, though, a picture is worth a thousand words.³⁰⁰ Until recently, design patent law had a rather straightforward effort to close the scope gap that did not require an explicit written scope definition: courts would line up the patented design, the accused product, and the closest piece of prior art next to each other.³⁰¹ Showing infringement of a valid design patent required persuading a single factfinder that the defendant's product was more like the patent than it was like the prior art. If it was more like the prior art, the patentee was improperly extending her right by suing the defendant. This simple artifice allowed a factfinder to easily see efforts to overclaim the scope of a design patent right in ways that might tread on the prior art.³⁰² It wasn't perfect, because visual images left to "speak for themselves" can sometimes mislead.³⁰³ Thus, we may sometimes want to require the parties to articulate what is different about the visual works the court is comparing in order to make sure that the asserted novelty of the plaintiff's work (or the asserted similarity of the defendant's) is in fact something we want the law to protect. Still, having both the accused product and the prior art together in a room was a good start.³⁰⁴

300. Cf. Rebecca Tushnet, *Worth a Thousand Words: The Images of Copyright*, 125 HARV. L. REV. 683, 703-04 (2012).

301. See, e.g., *FMC Corp. v. Hennessy Indus., Inc.*, 836 F.2d 521, 527 (Fed. Cir. 1987); *Applied Arts Corp. v. Grand Rapids Metalcraft Corp.*, 67 F.2d 428, 430 (6th Cir. 1933).

302. See Rebecca Tushnet, *The Eye Alone Is the Judge: Images and Design Patents*, 19 J. INTELL. PROP. L. 409, 411 (2012) ("Many designs involve non-novel or functional features, and are only protectable because of other novel features, or because of a novel combination of otherwise familiar features. But if the overall effect of the design on the eye is the measure of infringement, how are factfinders to avoid potentially broadening the scope of the design patent too far when an accused design looks like a patented design because of similarities in these unprotectable features?").

303. For discussions of the difficulties courts have in evaluating nonwritten works in IP, see, for example, Jessica Silbey, *Evidence Verité and the Law of Film*, 31 CARDOZO L. REV. 1257, 1280-81 (2010); Jessica M. Silbey, *Judges as Film Critics: New Approaches to Filmic Evidence*, 37 U. MICH. J.L. REFORM 493, 508-09, 554-57 (2004); Tushnet, *supra* note 300, at 696, 697 & n.94. As we noted above, the "Blurred Lines" copyright case may be an example of this.

304. We acknowledge that this approach was especially challenging in cases in which the claimed design incorporated features of several prior designs, because a court then couldn't simply identify a single piece of prior art as the "closest." But we think the instinct was right,

Unfortunately, the Federal Circuit abolished this test in 2009, replacing it with an “ordinary observer” test that further separated validity from infringement and asked the factfinder deciding infringement to compare only the defendant’s product and the patented design. In theory this comparison is done by an ordinary observer “familiar with the prior art”—meaning that comparisons between the claimed design and the prior art would still be relevant under the ordinary observer test.³⁰⁵ But it is now up to the defendant to frame any such comparisons, as the factfinder will not have the benefit of the court identifying particularly the features that differentiate the claimed design from that prior art. In the *Prestige Jewelry* case, for example, the district court refused to characterize the patented design at all, suggesting that “[a] jury can count the number of peripheral diamonds and observe that the peripheral diamonds are smaller than the center diamond. A jury also can easily observe that the peripheral diamonds have a different cut from the center diamond.”³⁰⁶ In fact, the court was concerned that “[t]ranslating the ‘132 Patent’s drawing into words would not be helpful to the jury and might lead a jury to fixate too much on specific elements such as cut and count”—even though it was clear that cut and count were precisely the things that differentiated the patented design from the prior art.³⁰⁷

and that having even a few pieces of prior art as part of the comparison would often have been manageable.

305. *Egyptian Goddess, Inc. v. Swisa, Inc.*, 543 F.3d 665, 678 (Fed. Cir. 2008) (“Our rejection of the point of novelty test does not mean, of course, that the differences between the claimed design and prior art designs are irrelevant. To the contrary, examining the novel features of the claimed design can be an important component of the comparison of the claimed design with the accused design and the prior art. But the comparison of the designs, including the examination of any novel features, must be conducted as part of the ordinary observer test, not as part of a separate test focusing on particular points of novelty that are designated only in the course of litigation.”).

306. *Prestige Jewelry Int’l, Inc. v. BK Jewellery HK*, No. 1:11-cv-02930-LAP-DCF, at *30 (S.D.N.Y. Oct. 15, 2015) (Sept. 15, 2014 order denying summary judgment).

307. *Id.* In *International Seaway Trading Corp. v. Walgreens Corp.*, the Federal Circuit extended the ordinary observer test (and its rejection of the point of novelty) to the anticipation context. *See* 589 F.3d 1233, 1243 (Fed. Cir. 2009). The court said that novelty would be assessed by comparing the claimed invention to the allegedly-anticipating prior art reference from the perspective of an ordinary observer familiar with the prior art (in that context, art prior to the prior art reference). *See id.* at 1238-39. Yet without even mentioning a single reference that predated the prior art reference, the court had no trouble concluding that certain elements of the claimed design were “distinctly different” from the prior art reference, and other differences were “minor or trivial.” *Id.* at 1242-43.

The fact that design patent law still contemplates a three-way comparison makes it superior to the kind of acontextual comparisons that are more common in copyright and trademark law.³⁰⁸ But we think the more explicit, side-by-side comparison of the patent, the accused product, and the prior art that was common before *Egyptian Goddess* offered courts a chance to focus on scope without an explicit definition of what is protectable, a desirable feature of the law that is unfortunately now much harder to achieve. In other circumstances, expert evidence may be desirable or even necessary to resolve the scope issues, as is often true today in patent cases, as well as software and music copyright cases.

An integrated scope proceeding will allow courts to catch cases of overclaiming, preventing IP owners from extending their rights to things those rights were never meant to cover.³⁰⁹ At the same time, an integrated scope proceeding should enable courts to reduce or eliminate the “busted hand” form of invalidity. Courts can define the scope of an IP right in ways that cabin it appropriately. Most of the time, the result will be to preserve the validity of that narrowed legal right.³¹⁰

An example of how that might work comes from a 2015 design patent case in the Federal Circuit, *Ethicon Endo-Surgery, Inc. v. Covidien, Inc.*³¹¹ The design patents there involved surgical devices, and covered “the overall appearance of the ornamental design of the ‘U’-shaped trigger and the particular ornamental design of a rounded and fluted torque knob positioned above and forward from the

308. See *supra* Part I.A.

309. In Christi Guerrini’s words, a scope proceeding will increase the “faithfulness” of the legal claim to the proper scope of the IP right. Christi J. Guerrini, *Defining Patent Quality*, 82 *FORDHAM L. REV.* 3091, 3131-33 (2014).

310. One important caveat is that sometimes the novel element of a work lies not in any given component, but in a novel combination of components that are each individually known before. See Lemley, *supra* note 24, at 1257, 1270, 1275-76. Patent law has dealt with this problem in obviousness cases in which the invention is a combination of two or more things that are individually described in the prior art. See *id.* at 1265. Patent courts are careful to avoid assuming in hindsight that it would have been obvious to combine two different things. See *id.* at 1254. Copyright courts similarly sometimes worry about losing what is creative about a work by overdissecting it. See *supra* note 299 and accompanying text. We agree that combinations of known elements are often creative and worthy of protection. That said, courts worry too much about the inability to detect those combinations. If the plaintiff asserts that it is not any individual element but a novel combination she has contributed, courts in a scope proceeding can evaluate that claim on its merits. See generally Lemley, *supra* note 24.

311. 796 F.3d 1312 (Fed. Cir. 2015).

trigger.”³¹² One might imagine that the shape of a trigger for a medical device was functional, and indeed the district court so held.³¹³ The Federal Circuit reversed, noting that its test for functionality in design patents is “stringent” and concluding that because there were other possible shapes a trigger could take—albeit ones surgeons said were not as good—the design was not functional as a matter of law.³¹⁴

So far, so bad—and consistent with the narrowing of functionality in other cases.³¹⁵ But then the Federal Circuit did something important: it focused on the scope of the resulting patent.³¹⁶ The district court, it said, found functionality because it read the design patent too abstractly, to cover U-shaped triggers and fluted knobs:

Instead of focusing on whether the specific patented designs had a functional purpose—the continuously curved “U” shape of the open trigger having tapered handles with ends flaring outwards, the football-shape of the activation button, and the asymmetrically-fluted torque knob with a flat front face—the district court focused its *PHG* analysis on the functional characteristics that *any* design of an open trigger, button, and torque knob would have for the underlying ultrasonic shears.³¹⁷

Thus, while the specific designs were nonfunctional, the functionality of the elements conceived more broadly could still limit the scope of the patent:

If the overall appearance of a claimed design is not primarily functional, the design claim is not invalid, even if certain elements have functional purposes. The scope of that claim, however, must be limited to the ornamental aspects of the design, and does not extend to “the broader general design concept.”³¹⁸

312. *Id.* at 1327.

313. *See id.* at 1327-28.

314. *Id.* at 1328; *see id.* at 1330-32.

315. *See supra* notes 245-50 and accompanying text.

316. *See Ethicon*, 796 F.3d at 1333.

317. *Id.* at 1331.

318. *Id.* at 1333 (citation omitted) (quoting *OddzOn Prods., Inc. v. Just Toys, Inc.*, 122 F.3d 1396, 1405 (Fed. Cir. 1997)).

The result was a narrowed patent that the defendant's product did not infringe, because its similarities were all at a more abstract, unprotectable level:

On a general conceptual level, both designs include an open trigger, a small activation button, and a fluted torque knob in relatively similar positions within the underlying ultrasonic device. Similarity at this conceptual level, however, is not sufficient to demonstrate infringement of the claimed designs.... [B]ecause each of these components has a functional aspect, the underlying elements must be excluded from the scope of the design claims at this general conceptual level. And when the remaining ornamental features of those components are compared, as a whole, to the corresponding ornamental features of Covidien's accused ultrasonic surgical shears, the dissimilarities between the designs are plain.³¹⁹

We think the Federal Circuit approached the infringement part of *Ethicon* in exactly the right way. The court focused explicitly on the scope of the patent and made sure not to let the patentee's claim extend to things that were not legally protectable because they were functional.³²⁰ In doing so, it reached the right result in the case before it (a ruling for defendant) without concluding that there was nothing at all protectable in the plaintiff's design. The court could do that only because it considered the two issues together, around the fulcrum of claim construction.

True, some IP rights have no valid scope; they are nothing more than efforts to claim ownership of something actually created by someone else or properly dedicated to the public.³²¹ But outright invalidity of an IP right is relatively uncommon outside of patent law.³²² For instance, a copyright owner has usually created something, and that something will commonly be deserving of some protection, even if not very much. Trademark law is somewhere in between.³²³ But determining what it is the owner actually created

319. *Id.* at 1336.

320. *See id.* at 1328-29, 1333, 1337.

321. *See generally supra* Part I.A.

322. *See supra* notes 78-87 and accompanying text.

323. Some marks are generic or functional; they are not entitled to any protection at all. *See, e.g.,* N.Y. Pizzeria, Inc. v. Syal, 56 F. Supp. 3d 875, 881-82 (S.D. Tex. 2014) (rejecting

requires more than just looking at the work; it requires an assessment of what the IP owner should be entitled to protect in that work.³²⁴

trademark claim to the taste of Italian food). But other marks are entitled to some protection, even if functionality or distinctiveness limit the proper scope of that protection. *See* McKenna, *supra* note 97, at 843 & n.89.

324. The district court held in *Sheldon v. Metro-Goldwyn Pictures Corp.*:

The Copyright Office does not, when a book is offered for copyright, study any prior art, as does the Patent Office when a patent is sought. It grants the copyright, thus putting the protection of the law not only over the copyrighted book as an entirety, but over the original content of the book. It is then left to the courts, if litigation ensues, to say what that original content is, and to define the zone in which the copyright owner is protected.

In defining that zone it always has to be determined: (1) Whether some part of the zone claimed is not a part of a common ground, the heritage of all mankind, usually referred to as the public domain; or (2) whether some of the infringement claimed is not of matter which is not protected by copyright for some other reason.

Naturally the plaintiff always seeks to widen his protected zone and the defendant to narrow it.

It follows that the approach of a court to the problem of the infringement of a play cannot be purely that of an ordinary playgoer, for such a playgoer presumably has not the opportunity to determine the limits of the protected zone by the principles above outlined.

7 F. Supp. 837, 843-44 (S.D.N.Y. 1934) (Woolsey, J.).

The Second Circuit reversed:

We are to remember that it makes no difference how far the play was anticipated by works in the public demesne which the plaintiffs did not use. The defendants appear not to recognize this, for they have filled the record with earlier instances of the same dramatic incidents and devices, as though, like a patent, a copyrighted work must be not only original, but new. That is not however the law as is obvious in the case of maps or compendia, where later works will necessarily be anticipated. At times, in discussing how much of the substance of a play the copyright protects, courts have indeed used language which seems to give countenance to the notion that, if a plot were old, it could not be copyrighted. But we understand by this no more than that in its broader outline a plot is never copyrightable, for it is plain beyond peradventure that anticipation as such cannot invalidate a copyright. Borrowed the work must indeed not be, for a plagiarist is not himself pro tanto an "author"; but if by some magic a man who had never known it were to compose anew Keats's Ode on a Grecian Urn, he would be an "author," and, if he copyrighted it, others might not copy that poem, though they might of course copy Keats's.... If the copyrighted work is therefore original, the public demesne is important only on the issue of infringement.

81 F.2d 49, 53-54 (2d Cir. 1936) (Hand, J.) (internal citations omitted).

With due respect to Judge Hand, however, Judge Woolsey seems to have it right. The relevant question is not merely whether there is a copyright but what the copyright covers. That question of scope cannot be simply dismissed as relevant only to infringement.

If the *Markman* experience is any indication, a scope proceeding may well resolve many of the disputes over infringement pretrial, particularly in copyright and design patent cases. Once we know what a patent does and doesn't cover, it is a rare case in which the operation of the defendant's product is in dispute. Copyright cases will have more disputes even after courts decide the scope of the work, because copyright incorporates some measure of market substitution as well as technical similarity into its infringement and defense analysis.³²⁵ Courts do this today in some software copyright cases, applying the Second Circuit's abstraction-filtration-comparison test.³²⁶ That approach could profitably be expanded to many other types of copyright cases. Trademark law is more complicated because it depends so heavily on the consumer confusion test, but even there, identifying unprotectable elements, particularly in trade dress cases, may make it sufficiently clear that confusion is unlikely and that summary judgment would be appropriate.³²⁷

Scope proceedings may have another benefit in an important class of cases—those that assert overlapping IP rights. IP regimes have a variety of doctrines that are designed to channel protection into one or another regime. One goal of functionality in both copyright and trademark, for instance, is to force creators who want protection for the functional aspects of their design to obtain a utility patent, with its higher threshold for protection and shorter term, rather than a “backdoor patent” through another IP regime.³²⁸ The idea-expression dichotomy does the same thing in software copyright.³²⁹ And courts have created an entire new doctrine devoted to

325. See Fromer & Lemley, *supra* note 25, at 1299-1301, 1300 n.263 (arguing that copyright's fair use doctrine generally requires some form of market substitution as well as technical similarity before finding infringement).

326. *Comput. Assocs. Int'l, Inc. v. Altai, Inc.*, 982 F.2d 693, 706 (2d Cir. 1992). The Tenth Circuit applies this test to all types of copyright cases, see *Gates Rubber Co. v. Bando Chem. Indus., Ltd.*, 9 F.3d 823, 834 (10th Cir. 1993), but it is an outlier.

327. See, e.g., Yvette Joy Liebesman, *Rethinking Trademark Functionality as a Question of Fact*, 15 NEV. L.J. 202, 205 (2014) (arguing that functionality should be treated as a legal conclusion, albeit one with factual underpinnings).

328. See, e.g., McKenna, *supra* note 97, at 847 & n.105; A. Samuel Oddi, *The Functions of “Functionality” in Trademark Law*, 22 HOUS. L. REV. 925, 927-28, 961 (1985) (quoting *In re Deister Concentrator Co.*, 289 F.2d 496, 504 (C.C.P.A. 1961)).

329. ROBERT P. MERGES ET AL., *INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE* 491 (6th ed. 2012) (“[T]he idea-expression dichotomy channels protection for functional works toward the patent system, which applies a relatively high threshold for protection (novelty

preventing IP owners from using trademark law to revive abandoned or expired copyrights.³³⁰ But those doctrines only work if the courts are aware of the possibility that one IP regime is being used to circumvent limits on a different regime. And they are unlikely to focus on the relationship between the two regimes if only one doctrine comes before them at a time, particularly if the doctrines designed to avoid overlap arise in different parts of the case. By focusing courts' attention on the nature and scope of the rights at issue, an integrated scope proceeding is likely to highlight areas of potential overlap and help the law channel IP claims into the proper doctrinal category.

D. Alternatives to a Scope Proceeding

Markman hearings are not costless, and neither would scope proceedings be in other IP regimes. Adding a hearing requires both party and judicial resources. We think that cost would be worth it if we could improve outcomes in IP cases. But we also think that, on balance, a scope hearing might reduce the cost of IP litigation. First, a scope hearing will often make it possible to resolve a case on summary judgment by resolving the underlying disagreement that has led to litigation. That has been patent law's experience with *Markman* hearings.³³¹ Indeed, one proposal for patent reform currently pending in Congress is to hold earlier *Markman* hearings precisely because doing so will reduce the cost of patent litigation.³³² For the same reason, a scope hearing is likely to push many parties to settle their case before trial. And although a court hearing is costly, a jury trial is far more costly. Finally, a scope proceeding may reduce strategic behavior by both parties, reducing the number of meritless cases or defenses that are raised. So, in the long run, scope proceedings may prove cost-effective and lead to more accurate results.

and nonobviousness), requires examination by a skilled examiner, and affords protection for only 20 years.... Awarding protection for functional works through copyright law—with its low threshold for protection and much greater duration—would undermine the role of the patent system.”)

330. See *Dastar Corp. v. Twentieth Century Fox Film Corp.*, 539 U.S. 23, 31-32 (2003); Mark P. McKenna, *Dastar's Next Stand*, 19 J. INTELL. PROP. L. 357, 362 & n.15, 366 (2012).

331. See Allison & Lemley, *supra* note 277, at 977.

332. See generally Innovation Act of 2015, H.R. 9, 114th Cong. (2015).

Some kind of integrated mechanism for setting scope is thus likely a first-best alternative. But if no such mechanism is workable in copyright or trademark, then we think both copyright and trademark law need to rely more heavily on validity doctrines and more frequently refuse rights altogether when the scope of protection ought to be very narrow. As those doctrines are currently constituted, that solution is less than ideal, because works that are properly entitled to some protection may be denied protection altogether. But such an approach should have a salutary long-term effect, driving copyright and trademark owners, who in the first instance make claims about the scope of their rights, to aim for narrower rather than broader rights.

Another compromise might be to allow validity doctrines to be less than all-or-nothing defenses. If a defendant could prevail by showing that she was making functional use of a trademark or patented design, or that she was practicing the prior art in patent law, courts would implicitly be limiting the scope of the IP rights without having to make the all-or-nothing decision to invalidate the IP right altogether. One of us has suggested taking that approach in trademark's functionality doctrine, for instance.³³³ And some courts have taken small steps in that direction, only to have them reversed.³³⁴

A final compromise is to raise the bar for infringement. The doctrine of virtual identity in copyright law serves this purpose. The normal test for copyright infringement is "substantial similarity" between the two works.³³⁵ But in some cases, where the rights are "thin," courts worry that decisionmakers using the lay observer test to assess similarity will find infringement based on similarities in unprotectable rather than protectable elements of a work. Those courts have not invalidated the copyrights altogether under merger or originality, but they have raised the threshold for proving infringement, requiring that the defendant's work be "virtually identical" to the plaintiff's rather than only substantially similar to

333. See generally McKenna, *supra* note 97.

334. See, e.g., *Fleischer Studios, Inc. v. A.V.E.L.A., Inc.*, 636 F.3d 1115 (9th Cir. 2011), *rev'd on reh'g*, 654 F.3d 958 (9th Cir. 2011); *Rosetta Stone Ltd. v. Google, Inc.*, 730 F. Supp. 2d 531 (E.D. Va. 2010), *rev'd in relevant part*, 676 F.3d 144 (4th Cir. 2012).

335. As one of us has noted elsewhere, the term "substantial similarity" is unhelpful because it is itself used to mean two different things in copyright law. See Lemley, *supra* note 112, at 720-21.

it.³³⁶ The virtual identity standard offers a middle ground between invalidation and overclaiming, narrowing the effective scope of the copyright in circumstances in which overclaiming seems likely because of the limited creativity associated with the work. Virtual identity seems a logical test to apply to a variety of works in which creativity is highly constrained, from computer software to photography. Some have even suggested that we could do something similar for trademarks, providing narrow protection to “thin marks.”³³⁷ Unfortunately, the application of the doctrine has so far been uneven and relatively rare.

In short, we need more judicial guidance about the proper scope of IP rights. That guidance will ideally take the form of identifying cases of overclaiming and weeding them out pretrial. By contrast, relying on expert testimony at trial to resolve these issues does not strike us as a satisfactory solution. Not only is it more costly, but the experience of the IP system with “technical experts” hired to testify on each side of any given issue has not been encouraging. And experimental evidence has shown that juries are guided by their view of labor and free riding in deciding similarity even when told to ignore them.³³⁸

Even when a jury trial is necessary because critical facts are in dispute, courts can help guide jury decisions to avoid the gaps we identify by making clear what similarities cannot be the basis for a finding of infringement. At a bare minimum, courts can use threshold burdens like the virtual identity standard to try to weed out cases in which overclaiming is particularly likely. And if none of these approaches work, courts will have to be more willing to invalidate IP rights for overclaiming than they have so far proven to be.

336. See, e.g., *Apple Comput., Inc. v. Microsoft Corp.*, 35 F.3d 1435, 1443 (9th Cir. 1994); *Hoehling v. Universal City Studios, Inc.*, 618 F.2d 972, 974, 977, 979-80 (2d Cir. 1980).

337. Michael Grynberg, *Thick Marks, Thin Marks* (Dec. 17, 2015) (unpublished manuscript), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705082 [<https://perma.cc/GQ4K-XEJM>].

338. See Balganesch et al., *supra* note 155, at 287-88.

CONCLUSION

Scope is *the* critical question that underlies all IP regimes. Surprisingly, it is almost always implicit, not explicit. Because courts have not paid careful attention to the scope of IP rights, they have allowed the parties to manipulate the scope of their IP rights in socially undesirable ways. A more explicit focus on the breadth of the IP right will lead to better decisions, reducing both the endemic overclaiming of IP rights and the occasional backlash that invalidates those rights altogether.

