

REFRAMING SIMILARITY ANALYSIS IN COPYRIGHT

KEVIN J. HICKEY*

ABSTRACT

Copyright law lacks a coherent method to determine non-literal infringement. The core inquiry, “substantial similarity,” purports to assess whether two works are so alike that an accused work infringes the original. Substantial similarity is a fundamental limit on the scope of copyright, but it is plagued by confusion and governed by a series of arcane tests that differ in each circuit. Even more troubling, courts lack a consistent method to go about comparing two works and how the comparison between two works is framed. There is no consensus, for example, on whether the original work or the accused work should be used as the baseline when assessing similarity. Courts sometimes adopt the perspective of the original creator, and sometimes of the alleged infringer, in determining whether seemingly copyrightable expression has become an uncopyrightable idea or functional standard. Courts are even confused as to whether dissimilarities or new material added by the defendant have any relevance to the comparison.

This Article seeks to bring analytical clarity to copyright’s similarity analysis, with a focus on these often-implicit framing issues. It argues that how courts frame the comparison, more than the legal test applied, is strongly associated with case outcomes. It urges courts to take a consistent approach to framing issues in similarity analysis so as not to improperly bias the comparison in favor of either party. In particular, courts should adopt a flexible, contextual approach to framing. This method considers both the perspective of the original creator and of the alleged infringer, as relevant, in drawing the line between permissible and substantial copying. It rejects the rigid approach that predominates in the case law, and endeavors to consider all relevant information about what was copied, how it was used in context, and why. The result is a similarity

* Microsoft Fellow, Berkeley Center for Law and Technology, University of California, Berkeley, School of Law. J.D. (2008), NYU School of Law; A.B. (2005), Mathematics, Brown University. I thank Pamela Samuelson, Kathy Strandburg, Jeanne Fromer, Kathryn Hashimoto, Peter Menell, Patrick Gould, Katharine Moore, Michael Wolfe, Barton Beebe, and Barry Friedman for their comments and suggestions on earlier drafts. I am grateful to the Microsoft Corporation and the Berkeley Center for Law and Technology for their financial support.

analysis that is not only more consistent, but a robust and vital limitation on the scope of copyright.

TABLE OF CONTENTS

INTRODUCTION.....	682
I. SUBSTANTIAL SIMILARITY BASICS.....	688
A. <i>Elements of Copyright Infringement</i>	689
B. <i>Questionable Standards: The Various “Tests” for Substantial Similarity</i>	690
1. <i>The Ordinary Observer Test</i>	690
2. <i>The Extrinsic/Intrinsic Test</i>	692
3. <i>“Total Concept and Feel”</i>	693
4. <i>The Abstraction/Filtration/Comparison Test</i>	694
C. <i>Common Ground</i>	695
1. <i>The Idea-Expression Distinction</i>	695
2. <i>Merger, Functionality, and Scènes à Faire</i>	696
3. <i>Thin and Thick Protection</i>	697
D. <i>Do the Tests Matter?</i>	698
II. FRAMING PROBLEMS IN SUBSTANTIAL SIMILARITY	699
A. <i>The Timing Problem</i>	701
B. <i>The Baseline Problem</i>	707
C. <i>The Relevance of Dissimilarity</i>	713
III. A PROPOSED CONTEXTUAL APPROACH.....	718
A. <i>Applying the Idea-Expression Distinction Dynamically</i>	720
B. <i>A Qualitative, Contextual Approach to Substantiality and Dissimilarities</i>	725
CONCLUSION.....	731

INTRODUCTION

It is not too much to ask that copyright law have a coherent approach to copyright infringement. But substantial similarity, copyright law’s core infringement inquiry, is a mess. Once the law allows that non-exact copies are actionable, courts need some method to determine when two works are so alike that one should be deemed an actionable infringement of the other. To use a classic example, would *West Side Story* infringe *Romeo & Juliet* were the latter still under copyright? Or, to take a contemporary

dispute, is Robin Thicke's hit "Blurred Lines" too similar to Marvin Gaye's classic "Got to Give It Up"?¹ This question, which copyright calls "substantial similarity," is a famously elusive problem.² Though substantial similarity is acknowledged to be a fundamental limit on the scope of copyright, equal in importance to fair use, courts' attempt to craft a predictable, consistent similarity doctrine is widely considered an utter failure.³

The most obvious aspect of this failure is the long-standing circuit split over the appropriate legal standard for substantial similarity. At least a half dozen tests for similarity proliferate in the courts, depending on the jurisdiction.⁴ To make matters worse, these tests are unduly complex, even by the arcane standards of copyright law.⁵ For example, the dominant formulation in the Second Circuit looks to whether an "ordinary observer"—sometimes a "more discerning" observer—would regard the aesthetic appeal of the two works as the same.⁶ The Ninth Circuit has bifurcated similarity analysis into "extrinsic" and "intrinsic" steps, with the judge first dissecting and objectively comparing the work's elements, followed by a subjective comparison of similarity by the jury.⁷ Still

1. See *Williams v. Bridgeport Music, Inc.*, No. LA CV13-06004 JAK (AGRx), 2015 WL 4479500 (C.D. Cal. July 14, 2015); Ben Sisario & Noah Smith, *'Blurred Lines' Infringed on Marvin Gaye Copyright, Jury Rules*, N.Y. TIMES (Mar. 10, 2015), <http://www.nytimes.com/2015/03/11/business/media/blurred-lines-infringed-on-marvin-gaye-copyright-jury-rules.html>.

2. See BENJAMIN KAPLAN, AN UNHURRIED VIEW OF COPYRIGHT 48 (Columbia Univ. Press eds., 1967) ("We are in a viscid quandary once we admit that [copyrightable] 'expression' can consist of anything not close aboard the particular collocation in its sequential order."); Shyamkrishna Balganeshe et al., *Judging Similarity*, 100 IOWA L. REV. 267, 268 (2014) ("In practice, though, the complexities of the fair use doctrine pale in comparison to what is central to almost all cases of copyright infringement: the question of 'substantial similarity.'"); Katherine Lippman, *The Beginning of the End: Preliminary Results of an Empirical Study of Copyright Substantial Similarity Opinions in the U.S. Circuit Courts*, 2013 MICH. ST. L. REV. 513, 515 (alteration in original) (quoting ROBERT C. OSTERBERG & ERIC C. OSTERBERG, SUBSTANTIAL SIMILARITY IN COPYRIGHT LAW xxi (2012)) ("[T]he concept of substantial similarity—a 'sine qua non of every [copyright infringement] determination'—remains one of the most elusive in copyright law.")

3. See, e.g., Mark A. Lemley, *Our Bizarre System for Proving Copyright Infringement*, 57 J. COPYRIGHT SOC'Y U.S.A. 719, 719 (2010) (characterizing the predominant similarity tests as "bizarre" and "mak[ing] no sense"); Pamela Samuelson, *A Fresh Look at Tests for Nonliteral Copyright Infringement*, 107 NW. U. L. REV. 1821, 1821 (2013) (describing the "case law to be frustratingly obscure, ambiguous, and confusing").

4. See *infra* Part I.B.

5. See, e.g., 3 WILLIAM F. PATRY, PATRY ON COPYRIGHT § 9:94 (2016) (criticizing the abstraction-filtration-comparison test as "the most complicated copyright 'test' ever conceived").

6. *Boisson v. Banian, Ltd.*, 273 F.3d 262, 272–73 (2d Cir. 2001).

7. *Shaw v. Lindheim*, 919 F.2d 1353, 1356–57 (9th Cir. 1990); *Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp.*, 562 F.2d 1157, 1163–64 (9th Cir. 1977), *superseded on other grounds by statute*, 17 U.S.C. § 504(b), *as recognized in Segal v. Rogue Pictures*, 544 F. App'x 769 (9th Cir. 2013).

another strain of jurisprudence declines to compare the individual elements of the two works at all, looking instead to similarity in the “total concept and feel” of the works.⁸ A final approach is the specialized “abstraction, filtration, comparison” test, usually (but not always) applied to computer software, which first lays out the various levels of abstraction in the work, then filters out unprotected ideas and functional elements, and finally compares the remainder of the two works.⁹

The issue of which of these tests is best has received a fair bit of attention. Much commentary on substantial similarity defends one of the existing tests as superior or proposes to replace the confused doctrine with a new standard.¹⁰ Recent empirical evidence suggests, however, that this long-running debate may be beside the point as a practical matter: in terms of case outcomes, the particular test that a court uses does not appear to make much difference.¹¹ Instead, it may be that factfinders reach an intuitive conclusion about similarity when comparing the two works, and only then use the various legal formulations to provide an *ex post* rationale for their decisions.¹²

8. See, e.g., *Atkins v. Fischer*, 331 F.3d 988, 993 (D.C. Cir. 2003); *Roulo v. Russ Berrie & Co.*, 886 F.2d 931, 939 (7th Cir. 1989).

9. See, e.g., *Gates Rubber Co. v. Bando Chem. Indus., Ltd.*, 9 F.3d 823, 834 (10th Cir. 1993); *Computer Assocs. Int'l, Inc. v. Altai, Inc.*, 982 F.2d 693, 706–11 (2d Cir. 1992).

10. See, e.g., Jeannette Rene Busek, *Copyright Infringement: A Proposal for a New Standard for Substantial Similarity Based on the Degree of Possible Expressive Variation*, 45 UCLA L. REV. 1777, 1795–1803 (1998) (proposing substantial similarity test tailored to the work's medium); Laura G. Lape, *The Metaphysics of the Law: Bringing Substantial Similarity Down to Earth*, 98 DICK. L. REV. 181, 194–206 (1993) (proposing harm-based approach to substantial similarity); Lawrence Jeffrey Sher, Comment, *The Search for a Suitable Standard of Substantial Similarity: The Ninth Circuit's Application of the Krofft Test*, 25 U.C. DAVIS L. REV. 229, 254–62 (1991) (defending variation on the *Krofft* test); Carl A. Sundholm, Essay, *High Technology Jurisprudence: In Defense of “Look and Feel” Approaches to Copyright Protection*, 8 SANTA CLARA COMPUTER & HIGH TECH. L.J. 209, 217–25 (1992) (defending the holistic “look and feel” approach); Daniel Su, Note, *Substantial Similarity and Architectural Works: Filtering Out “Total Concept and Feel,”* 101 NW. U. L. REV. 1851, 1878–83 (2007) (arguing for wider use of the abstraction, filtration, and comparison test).

11. See Lippman, *supra* note 2, at 545 (“Remarkably, however, when collapsing the six substantial similarity tests into the three main tests—the ordinary observer test, the extrinsic/intrinsic test, and the abstraction-filtration-comparison test—their substantial similarity win rates become closely aligned.”).

12. See, e.g., Zahr K. Said, *Reforming Copyright Interpretation*, 28 HARV. J.L. & TECH. 469, 473–75, 522 (2015) (noting a “great divergence in outcomes and reasoning . . . in infringement analysis” and criticizing judges’ “intuitive” approach); B. MacPaul Stanfield, Note, *Finding the Fact of Familiarity: Assessing Judicial Similarity Tests in Copyright Infringement Actions*, 49 DRAKE L. REV. 489, 512 (2001) (“Perhaps the [substantial similarity tests] are not a means to determine similarity, but rather a means to explain a finding of similarity that is determined in such a way that defies clear explanation.”). There is evidence of an analogous phenomenon in the “likelihood of confusion” test in trademark law, where a few factors seem to drive the decision, and judges use the other factors to merely justify the outcome. See Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581, 1598–1622 (2006) (describing trademark

Judicial confusion about substantial similarity has another, deeper dimension, however—one that has largely escaped notice. This concerns not the ultimate legal threshold for similarity, but instead how the comparison of the original work and the accused work is *framed*—that is, how courts go about comparing the two artistic works at issue. For example, when looking for similarities, should the factfinder assume the perspective of the original creator or of the alleged infringer? Results from the cognitive sciences have shown that such framing effects can influence whether or not people perceive two things as similar.¹³ These framing issues, then, may be more important to case outcomes than the much-debated legal tests. This Article seeks to turn attention to the often implicit aspects of *how* courts approach similarity analysis. Courts have not taken a consistent approach to these issues, to put it mildly.

Consider as an initial example what we will call similarity's *timing problem*. Much of similarity analysis depends upon distinguishing between the elements of a copyrighted work that are protectable "expression," as opposed to uncopyrightable "ideas."¹⁴ In a work of nonfiction, for example, the expression would tend to include things like particular word choices and phrasing, while the broader thesis and claims would be uncopyrightable ideas. Simply put, a later creator is free to advance the same argument in a new work, so long as she expresses the point in a different way. Ideas, however, are not static concepts. For example, the QWERTY keyboard might have been an original arrangement of letters when it was created, but it is now surely a functional standard idea that cannot be copyrighted.¹⁵ This raises the question: at what point in time should similarity be assessed? Should we apply the idea-expression distinction at the time of an original work's creation, or at the time of alleged copyright infringement? Courts have not settled on an answer.¹⁶

"decision making in which certain factors drive the outcome and the rest of the factors subsequently fall in line to support that outcome").

13. See generally Kate Klonick, *Comparing Apples to Applejacks: Cognitive Science Concepts of Similarity Judgment and Derivative Works*, 60 J. COPYRIGHT SOC'Y U.S.A. 365, 383 (2013) ("Framing can result in asymmetrical similarity judgments depending on what is being considered as the referent and what is being considered as the subject."); Amos Tversky, *Features of Similarity*, 84 PSYCHOL. REV. 327, 340 (1977) ("Like other judgments, similarity depends on context and frame of reference.").

14. 17 U.S.C. § 102(b) (2014). In addition to "ideas," the statute also prohibits copyright protection for a "procedure, process, system, method of operation, concept, principle, or discovery." *Id.*

15. See *Lotus Dev. Corp. v. Borland Int'l, Inc.*, 49 F.3d 807, 819–20 (1st Cir. 1995) (Boudin, J., concurring) (discussing the QWERTY keyboard example and noting that copyright protection in that case would "present the concern with fencing off access to the commons in an acute form").

16. See, e.g., *Lexmark Int'l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522, 557 (6th Cir.

Similarity's *baseline problem* is another divide relating to how the infringement analysis is framed.¹⁷ All courts agree that copying must be "substantial" to be actionable, but substantial relative to what? Naively, one might presume to look at the amount of the original work that was appropriated, or to the amount of the allegedly infringing work that was copied. But these measures are not the same if the original work is larger than the accused work, or vice versa. Is copying a page from an encyclopedia less wrongful than copying the same quantity of expression from a novel? Does it matter whether the infringer incorporated that page into a larger work? In practice, courts sometimes reject using the defendants' work as the baseline, for fear that a plagiarist could "excuse the wrong by showing how much of his work he did not pirate."¹⁸ But the alternative of using the plaintiff's work as the measure is no more equitable.¹⁹ In effect, that baseline penalizes the original author for creating more expression, which seems perverse if copyright is intended to incentivize the creation of new works.²⁰ Regardless of which baseline is best, it might be hoped that courts would at least be consistent, but they have reached no consensus.²¹ Even the Supreme Court, addressing the issue in the related context of fair use, switched between various baselines in its opinion without acknowledging the inconsistency.²²

A third framing aspect of substantial similarity—the *dissimilarity problem*—relates to the weight, if any, that dissimilarities between the two works should have in the analysis. In other words, if two works share a substantial number of similarities, should the fact that they differ in other

2004) (Feikens, J., concurring in part and dissenting in part) (reviewing split in authority as to whether idea-expression merger is assessed at copyrightability stage or infringement stage); *infra* Part II.A.

17. See *infra* Part II.B.

18. *Sheldon v. Metro-Goldwyn Pictures Corp.*, 81 F.2d 49, 56 (2d Cir. 1936).

19. *Contra* 3 PATRY, *supra* note 5, § 9:65 (collecting cases and asserting that "the only relevant inquiry is whether what was copied from the plaintiff is a material part of the plaintiff's work").

20. See *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 558 (1985) ("[T]he Framers intended copyright itself to be the engine of free expression. By establishing a marketable right to the use of one's expression, copyright supplies the economic incentive to create and disseminate ideas.").

21. Compare *Peter Letterese & Assocs. v. World Inst. of Scientology Enters.*, 533 F.3d 1287, 1307 (11th Cir. 2008) ("[I]t is the relative portion of the copyrighted work—not the relative portion of the infringing work—that is the relevant comparison."), with *Narell v. Freeman*, 872 F.2d 907, 913 (9th Cir. 1989) (emphasis added) ("[I]nsignificant infringement may be substantial only if the material is qualitatively important to either [the accused or the original] work.").

22. See *Harper & Row*, 471 U.S. at 564–66 (using multiple baselines in noting that amount copied was a quantitatively "insubstantial" part of the original work but also "13% of the infringing article").

ways matter?²³ At least one leading commentator thinks not, stating categorically, “[i]t is entirely immaterial that, in many respects, plaintiff’s and defendant’s works are dissimilar, if in other respects, similarity as to a substantial element of plaintiff’s work can be shown.”²⁴ The case law, however, routinely considers “points of dissimilarity” as relevant and tending to undermine a finding of similarity²⁵—though not always.²⁶ Again, courts show distressing inconsistency on a seemingly fundamental aspect of how to go about comparing two works.

This Article attempts to bring some analytical clarity to substantial similarity doctrine, with a particular focus on the framing problems that have largely avoided scholarly attention. Its principal aim is to examine the overlooked and often-implicit ways that courts frame their comparisons, and to call for a consistent approach to these conceptual problems in similarity analysis. Detailed examination of several examples suggests that these aspects of substantial similarity may be more important to the outcome than the particular legal test that is used.²⁷ After analyzing the widespread inconsistency of the courts on these issues, the Article offers some proposed resolutions based on copyright’s core objective to “promote the Progress of Science.”²⁸ Whether or not the reader ultimately agrees with those resolutions is not the primary point, though of course I do hope to persuade. More important, in my view, is for courts to become aware of the implicit biases that framing can create in decision-making, and to adopt a consistent, principled approach to framing in similarity

23. Clearly, the addition of new material has relevance for the fair use inquiry. *See Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994) (considering whether the accused work “adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message” under the first fair use factor). Whether new or transformative material is also relevant to the question of substantial similarity, however, is an independent and unresolved issue. *See infra* Part II.C.

24. 4 MELVILLE B. NIMMER, NIMMER ON COPYRIGHT § 13.03[B][1][a] (David Nimmer rev. ed., 2014).

25. *See, e.g., Zaleski v. T.P. Builders, Inc.*, 875 F. Supp. 2d 135, 148 (N.D.N.Y. 2012) (citations omitted) (“[D]issimilarity can be important in determining whether there is substantial similarity. . . . [N]umerous differences tend to undercut substantial similarity.”).

26. *See, e.g., Attia v. Soc’y of the N.Y. Hosp.*, 201 F.3d 50, 57 (2d Cir. 1999) (quoting *Novelty Textile Mills, Inc. v. Joan Fabrics Corp.*, 558 F.2d 1090, 1093 n.4 (2d Cir. 1977)) (“[T]he key [to substantial similarity is] the similarities rather than the differences.”); *Concrete Mach. Co. v. Classic Lawn Ornaments, Inc.*, 843 F.2d 600, 608 (1st Cir. 1988) (“At times, the existence of only minor differences may itself suggest copying, indicating that the infringer attempted to avoid liability by contributing only trivial variations.”).

27. *See infra* Part II; *cf. Lippman, supra* note 2, at 545–46 (finding that as to the three main substantial similarity tests, “the probability that the copyright holder will prevail remains the same” regardless of which test is used).

28. U.S. CONST. art. I, § 8, cl. 8.

analysis. The current confusion frees courts to frame the comparison in whatever way best justifies a desired outcome. The resulting unpredictability ill serves courts, litigants, and—critically—primary actors attempting to conform their behavior to the law.

The remainder of this Article will proceed as follows. Part I will review the confused state of substantial similarity doctrine, including both the various tests and elements, like the idea-expression distinction, that are common across all jurisdictions. Part II will discuss the divisions in how courts frame the comparison between two works, presenting several case law examples that illustrate the confusion surrounding these aspects of similarity analysis. Finally, Part III will analyze the conflicts over framing in substantial similarity and offer resolutions that will serve to promote judicial consistency, predictability, and copyright's ultimate goal of encouraging the creation and dissemination of knowledge.

I. SUBSTANTIAL SIMILARITY BASICS

This Part will provide a brief overview of the current state of substantial similarity doctrine. The first Part lays out the elements of copyright infringement in order to understand how substantial similarity fits into a typical case. The next Part briefly reviews the various legal standards for substantial similarity as they have developed in different circuit courts of appeals. The third Part turns to those elements of similarity analysis that are common across all jurisdictions, regardless of the particular test used. The final Part steps back to ask whether these varied and confusing standards have as much practical import as is sometimes claimed.

Two themes emerge from this sketch of the state of substantial similarity. The first is the unusual diversity of tests. Unlike the typical “circuit split,” the courts are not simply lined up into two opposing camps; there are six or more competing formulations, depending on the classification scheme.²⁹ Moreover, even within a single circuit, courts sometimes vary the test used depending on the type of work at issue (e.g.,

29. See generally ROBERT C. OSTERBERG & ERIC C. OSTERBERG, SUBSTANTIAL SIMILARITY IN COPYRIGHT LAW § 3 (2015) (reviewing each circuit's tests for substantial similarity, including the ordinary observer, extrinsic/intrinsic, abstraction/filtration/comparison, and intended audience tests) [hereinafter OSTERBERG]; Lippman, *supra* note 2, at 546 (comparing success rate of six different tests for substantial similarity); Samuelson, *supra* note 3, at 1823–37 (reviewing five conventional substantial similarity tests including the ordinary observer, extrinsic/intrinsic, abstractions, abstraction/filtration/comparison, and total concept and feel tests).

software vs. visual art).³⁰ The second theme is the roundabout ways in which these diverse tests, although formulated differently, tend to converge upon similar concepts.³¹

A. *Elements of Copyright Infringement*

At the broadest level, a successful plaintiff in a copyright infringement suit must prove just two things: (1) ownership of a valid copyright in a work, and (2) improper copying of original expression by the defendant.³² The second element can be further broken down into two aspects. The first (2a) is actual copying: proof that the defendant, as a matter of fact, copied from the plaintiff's work as opposed to creating her work independently.³³ The second (2b) is that the copying of protected material was substantial enough to amount to an "improper appropriation."³⁴

It is critical to distinguish the role that similarity sometimes plays in proof of actual copying (element 2a) from the role that similarity plays in determining whether the copying was improper (element 2b). The former is sometimes called "probative similarity," and it relates only to the factual issue of proving that the defendant did not create her work on her own, but instead copied from the plaintiff's work.³⁵ Actual copying can be proven either directly (such as by admission of the defendant) or, more commonly, indirectly.³⁶ Indirect proof of copying usually consists of proving that the defendant had access to the original work, and that the similarities between the original and the accused work are suspicious enough to conclude that the defendant copied.³⁷ Similarity between the two works can thus be relevant to the question of actual copying, but it is *not* what is meant by the term "substantial similarity."³⁸

30. To take a common example, most circuits use the abstraction/filtration/comparison test in computer software cases, regardless of the test used for other literary works. *See* Samuelson, *supra* note 3, at 1837–40 (describing the abstraction/filtration/comparison test as "[t]he most widely used test for judging nonliteral infringement of computer programs").

31. *See infra* Part I.C.

32. *See* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 361 (1991); *Arnstein v. Porter*, 154 F.2d 464, 468 (2d Cir. 1946).

33. Proof of actual copying is necessary as independent creation negates a claim of copyright infringement. *See, e.g.,* Calhoun v. Lillenas Publ'g, 298 F.3d 1228, 1233 (11th Cir. 2002).

34. *Arnstein*, 154 F.2d at 468.

35. *See* Ringgold v. Black Entm't Television, Inc., 126 F.3d 70, 74–75 (2d Cir. 1997) (citing Alan Latman, "Probative Similarity" as Proof of Copying: Toward Dispelling Some Myths in Copyright Infringement, 90 COLUM. L. REV. 1187, 1204 (1990) (distinguishing between substantial similarity and probative similarity)).

36. *Arnstein*, 154 F.2d at 468.

37. *Id.*

38. Regrettably, even sophisticated courts sometimes confuse probative similarity with

Substantial similarity refers instead to the knottier question of when copying becomes so extensive that it should be legally actionable. Not all copying is copyright infringement. For example, copying might be permissible because what was copied was unoriginal, in the public domain, unprotected by copyright, or *de minimis*.³⁹ Substantial similarity, then, is used as a term of art to represent when copying of protected material is quantitatively and qualitatively significant enough to be an infringement of copyright.⁴⁰ To say that two works are substantially similar is just another way of saying that they are so alike in protected elements that one work infringes upon the other. It is the conclusion that the defendant's appropriation was improper. As the next Part reveals, courts have created an obscure array of legal tests in their attempt to draw this line.

B. Questionable Standards: The Various "Tests" for Substantial Similarity

This Part reviews the leading formulations for substantial similarity, and the difficulties with each approach. These tests are not the primary focus of this Article, and—as I argue below—they may have received an undue amount of attention in the existing commentary on substantial similarity. However, an understanding of the ultimate legal standards for substantial similarity is essential to appreciate the importance of the framing problems that are discussed in the next Part.

1. The Ordinary Observer Test

One of the earliest tests for substantial similarity—and still the most widely used—is the Second Circuit's "ordinary observer" test. The reason for the test's longevity may be its simplicity. It just asks whether a typical

substantial similarity. *See, e.g.,* Lemley, *supra* note 3, at 720–21 (criticizing judicial decisions for making this mistake).

39. *See* Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23, 33–34 (2003) (“[O]nce the patent or copyright monopoly has expired, the public may use the invention or work at will and without attribution.”); Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 547 (1985) (citing 17 U.S.C. § 102) (“[N]o author may copyright facts or ideas. . . . The copyright is limited to those aspects of the work—termed ‘expression’—that display the stamp of the author’s originality.”); Newton v. Diamond, 388 F.3d 1189, 1193 (9th Cir. 2003) (“[E]ven where the fact of copying is conceded, no legal consequences will follow from that fact unless the copying is substantial.”).

40. *See* Shyamkrishna Balganesh, *The Normativity of Copying in Copyright Law*, 62 DUKE L.J. 203, 205, 208 (2012) (characterizing substantial similarity as a “subjective evaluation of the different parts of the two works and of their relative contributions to the overall significance of the work, both as a quantitative and qualitative matter”).

lay observer would recognize the accused work as having been copied from the original.⁴¹ In other words, would a reasonable viewer, “unless he set out to detect the disparities [between the two works], . . . be disposed to overlook them, and regard their aesthetic appeal as the same”?⁴² The ordinary observer test derives from the influential case *Arnstein v. Porter*, in which the renowned composer Cole Porter was accused of plagiarizing the comparatively unknown works of Ira Arnstein.⁴³ The legal issue in *Arnstein* was whether the similarity analysis should be informed by “dissection” of the work (e.g., analysis of the score) and expert testimony, or just the music’s effect on a “lay listener.”⁴⁴ The court held that the latter was the key for substantial similarity, but that expert testimony was permitted to prove probative similarity.⁴⁵ Variations of the ordinary observer test are followed in the First, Third, Fifth, and Seventh Circuits.⁴⁶

The most obvious difficulty with the ordinary observer test is that, by itself, it provides “scant guidance” in similarity analysis.⁴⁷ While it specifies the audience that is to assess similarity, it does not specify what the lay observer should be looking for. Moreover, the ordinary observer test fails to account for the possibility that the perceived similarity between the two works arises primarily from uncopyrightable elements. Perhaps as a result, the Second Circuit has created a “more discerning observer” test that applies when the work at issue contains significant unprotected elements, such as material taken from the public domain.⁴⁸ A final shortcoming of the test is that, with regard to dissection and expert testimony, *Arnstein* gets the matter “exactly backwards.”⁴⁹ For probative similarity, dissection of the work is not necessary—copying of unprotectable elements is relevant to prove that the accused work was not independently created. But the legal conclusion of substantial similarity—that what was copied was original and significant enough to infringe—is precisely the type of issue where a factfinder may need guidance from expert musicologists, computer scientists, and the like.⁵⁰

41. *Knitwaves, Inc. v. Lollytogs Ltd.*, 71 F.3d 996, 1002 (2d Cir. 1995).

42. *Boisson v. Banian, Ltd.*, 273 F.3d 262, 272 (2d Cir. 2001) (quoting *Folio Impressions, Inc. v. Byer Cal.*, 937 F.2d 759, 765 (2d Cir. 1991)).

43. *Arnstein v. Porter*, 154 F.2d 464, 467 (2d Cir. 1946).

44. *Id.* at 468–69.

45. *Id.* at 473.

46. See OSTERBERG, *supra* note 29, §§ 3:1.2–.5; Lippman, *supra* note 2, at 526–27.

47. See Samuelson, *supra* note 3, at 1825.

48. *Boisson v. Banian, Ltd.*, 273 F.3d 262, 272 (2d Cir. 2001).

49. See Lemley, *supra* note 3, at 736–38.

50. See *id.*; Samuelson, *supra* note 3, at 1840–42.

2. *The Extrinsic/Intrinsic Test*

The extrinsic/intrinsic test was created by the Ninth Circuit in the 1970s and later adopted by the Fourth and Eighth Circuits.⁵¹ The leading case is *Sid & Marty Krofft Television Productions, Inc. v. McDonald's Corp.*, which involved whether the familiar McDonaldland characters (Ronald McDonald, Grimace, the Hamburglar et al.) infringed the characters of an earlier children's television program, *H. R. Pufnstuf*.⁵² In the course of affirming the judgment against McDonald's, the Ninth Circuit crafted a two-step test for substantial similarity. The first "extrinsic" step asks whether there is similarity in the ideas of the two works, focusing on objective criteria such as "the type of artwork involved, the materials used, the subject matter, and the setting"; at this stage, dissection of the work and expert testimony are permitted.⁵³ The second "intrinsic" step asks whether there is similarity in the expression of the works, focusing on the subjective response of the factfinder.⁵⁴ Some circuits refine this second step further by assuming the perspective of the "intended audience" of the work, as opposed to the general public.⁵⁵

Like the ordinary observer test, the extrinsic/intrinsic test suffers from several shortcomings. First, the extrinsic step's focus on similarity in *ideas* is completely wrongheaded, because copyright does not protect ideas.⁵⁶ Thankfully, the Ninth Circuit eventually abandoned this formulation, explaining that the extrinsic and intrinsic steps both analyze similarity in protected expression: it is simply that the former focuses on objective criteria, whereas the intrinsic step embraces a more subjective, impressionistic approach.⁵⁷ The extrinsic/intrinsic terminology is also needlessly complicated.⁵⁸ While "objective/subjective" might better

51. See *Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp.*, 562 F.2d 1157, 1163–64 (9th Cir. 1977), *superseded on other grounds by statute*, 17 U.S.C. § 504(b), *as recognized in Segal v. Rogue Pictures*, 544 F. App'x 769 (9th Cir. 2013); OSTERBERG, *supra* note 29, §§ 3:2.1–3; Lippman, *supra* note 2, at 529–30.

52. *Krofft*, 562 F.2d at 1161–62.

53. *Id.* at 1164.

54. *Id.*

55. See, e.g., *Dawson v. Hinshaw Music Inc.*, 905 F.2d 731, 733–36 (4th Cir. 1990).

56. 17 U.S.C. § 102(b) (2014) ("In no case does copyright protection for an original work of authorship extend to any idea . . .").

57. *Shaw v. Lindheim*, 919 F.2d 1353, 1356–57 (9th Cir. 1990).

58. See, e.g., NIMMER, *supra* note 24, § 13:03[E][3][b] (describing the extrinsic/intrinsic test as a "metaphysical palace[]"); 3 PATRY, *supra* note 5, § 9:235 (describing the extrinsic/intrinsic terminology as "infinitely . . . complicated" and "meaningless"); Samuelson, *supra* note 3, at 1829 (describing the extrinsic/intrinsic terminology as "inapt and confusing").

capture the distinction between the two steps, the old terms persist nonetheless.⁵⁹

3. “Total Concept and Feel”

The “total concept and feel” approach to similarity originated with another influential Ninth Circuit case, *Roth Greeting Cards v. United Card Co.*⁶⁰ In *Roth*, the defendant copied short phrases from Roth’s greeting cards, such as “I miss you already” (on the front of the card), and “You Haven’t even Left” (on the inside of the card).⁶¹ The defendant created its own drawings to accompany those simple messages, however. The district court in *Roth* found no infringement because the copied language was uncopyrightable (it was too short to be original), and the images on the cards were not similar.⁶² The Ninth Circuit reversed; it agreed that the language was not protectable, and that the images were different.⁶³ Nonetheless, it found sufficient similarity in the “total concept and feel” of the cards, such as “the combination of art work conveying a particular mood with a particular message.”⁶⁴ Judge Kilkenney, in a strong dissent, was perplexed at how the whole of the work—the “total feel” of the card—could be any “greater than the sum total of its parts”: the words plus the images.⁶⁵

From these humble beginnings, “total concept and feel” rose to become a common intonation in substantial similarity cases, either as a part of the other tests or as its own stand-alone test.⁶⁶ The phrase and its cousin, “overall look and feel,” are invoked on occasion by all the circuits and across a wide variety of subject matter.⁶⁷ The test’s ubiquity belies its glaring conceptual problems. Like the extrinsic step, “total concept and feel” wrongly assesses similarity in elements—“concepts”—that are not

59. See *Shaw*, 919 F.2d at 1357 (“[T]he two tests are more sensibly described as objective and subjective analyses of *expression* . . .”).

60. 429 F.2d 1106 (9th Cir. 1970), *superseded on other grounds by statute*, 17 U.S.C. § 411(a), as recognized in *Cosmetic Ideas, Inc. v. IAC/InteractiveCorp.*, 606 F.3d 612 (9th Cir. 2010).

61. *Id.* at 1110.

62. *Id.* at 1109.

63. See *id.* at 1109–10.

64. *Id.* at 1110.

65. *Id.* at 1111–12 (Kilkenney, J., dissenting).

66. See, e.g., *Atkins v. Fischer*, 331 F.3d 988, 993 (D.C. Cir. 2003); *Roulo v. Russ Berrie & Co.*, 886 F.2d 931, 939 (7th Cir. 1989).

67. See generally OSTERBERG, *supra* note 29, § 2:5.2 (collecting cases); Samuelson, *supra* note 3, at 1830–31 nn. 49–50 (same); NIMMER, *supra* note 24, § 13:03[A][1][c] (same).

protected by copyright.⁶⁸ Just as fundamentally, the test threatens an unduly broad notion of infringement, permitting a finding of substantial similarity even when the plaintiff cannot articulate any specific elements that were copied.⁶⁹

4. *The Abstraction/Filtration/Comparison Test*

The abstraction/filtration/comparison test (the “A/F/C test”) arose out of the difficulties of assessing similarity in computer software cases.⁷⁰ At its core, the A/F/C test is simply a formalized system to ensure that elements that are not protected by copyright are excluded when comparing two works. First, the abstraction step examines the plaintiff’s work to separate the general ideas—elements that are unoriginal, too general to be protected, dictated by functional concerns, and the like—from protectable, original expression.⁷¹ The second step filters out these unprotected elements. In the final step, only the protected expression of the two works is compared.⁷² Most circuits apply the A/F/C test in computer software cases, regardless of the test that they apply to other works.⁷³ In the Tenth Circuit, however, the A/F/C test is applied to *all* copyrighted works, and the Sixth Circuit likewise applies its variant of the A/F/C test across the board.⁷⁴

There is much to admire about the A/F/C test, particularly in its emphasis on distinguishing between protectable and unprotectable elements, and excluding the latter from similarity analysis. But the A/F/C test has also received its fair share of criticism. Former Professor Patry is perhaps the harshest critic, calling A/F/C “the most complicated copyright ‘test’ ever conceived” that misses important holistic elements of artistic works by the “brutal . . . desiccation” of element-by-element analysis.⁷⁵ Others argue that, whatever usefulness the A/F/C test may have in the

68. See 4 NIMMER, *supra* note 24, § 13:03[A][1][c] (footnote omitted) (citations omitted) (“More broadly, the touchstone of ‘total concept and feel’ threatens to subvert the very essence of copyright ‘Concepts’ are statutorily ineligible for copyright protection”).

69. See Su, *supra* note 10, at 1871–72 (noting that “works may appear substantially similar in their ‘total concept and feel’ even though the similarities arise from the common use of unprotectable elements, such as abstract ideas or public domain material,” and collecting examples in the architectural context).

70. See, e.g., *Computer Assocs. Int’l, Inc. v. Altai, Inc.*, 982 F.2d 693, 706–11 (2d Cir. 1992).

71. See, e.g., *Country Kids ’N City Slicks, Inc. v. Sheen*, 77 F.3d 1280, 1284–85 (10th Cir. 1996).

72. *Id.* at 1285.

73. See OSTERBERG, *supra* note 29, § 3; NIMMER, *supra* note 24, § 13:03[F] nn.283.6–9.

74. See OSTERBERG, *supra* note 29, § 3; Lippman, *supra* note 2, at 531–32.

75. 3 PATRY, *supra* note 5, § 9:94; 2 WILLIAM F. PATRY, PATRY ON COPYRIGHT § 4:27 (2016).

software context, it makes little sense for visual or musical works to be dissected in this manner.⁷⁶

C. Common Ground

Despite the startling diversity of tests for substantial similarity, there may be more common ground between the circuits than is usually appreciated. Regardless of the test used, courts apply similar concepts in distinguishing between unprotectable and protectable elements. This Part reviews those doctrines—most prominently the idea-expression distinction—that play a role in similarity analysis in all jurisdictions.

1. The Idea-Expression Distinction

Perhaps the most fundamental principle of copyright is the notion that only original expression, and not “ideas,” can be the subject of copyright. Section 102(b) of the Copyright Act states that “[i]n no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery.”⁷⁷ This codifies the holding of *Baker v. Selden*, an 1879 Supreme Court case addressing the extent of copyright in a book describing a method of accounting.⁷⁸ The Court held that the copyright extended only to the explanation in the book, and not to the method described therein: Selden could prevent others from copying the text of the book, but not from using or describing his accounting method.⁷⁹ Copyright is intended to promote the creation and dissemination of knowledge, and granting a monopoly over something as general as an idea could suppress the flow of knowledge.⁸⁰

While the intuition behind the idea-expression distinction is easy to grasp, precisely defining the line between idea and expression is not easy, and perhaps not possible.⁸¹ Nevertheless, the idea-expression distinction

76. See, e.g., Jon O. Newman, *New Lyrics for an Old Melody: The Idea/Expression Dichotomy in the Computer Age*, 17 *CARDOZO ARTS & ENT. L.J.* 691, 698 (1999) (“[O]ne cannot divide a visual work into neat layers of abstraction in precisely the same manner one could with a text.”).

77. 17 U.S.C. § 102(b) (2014).

78. *Baker v. Selden*, 101 U.S. 99, 100 (1879).

79. *Id.* at 102–04.

80. See *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 349–50 (1991) (citing *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556–57 (1985)) (“[C]opyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work.”).

81. The best attempt still belongs to Judge Learned Hand, who wrote when analyzing similarity between a play and a motion picture:

lurks behind much of substantial similarity analysis, either explicitly or implicitly. It is most apparent in the A/F/C test, which is directed primarily at extracting unprotectable “ideas” from the similarity analysis. But it is present in the observer test as well: the supposedly “ordinary” observer tends to become “more discerning” precisely when there are considerable unprotectable ideas at issue.⁸² In a similar manner, the extrinsic/intrinsic test, as modified, applies only to expression, not ideas.⁸³

2. *Merger, Functionality, and Scènes à Faire*

The doctrines of merger and *scènes à faire* can be viewed as corollaries of the idea-expression distinction, and they, too, are commonly invoked in similarity analysis regardless of the test being applied. The merger doctrine reasons that when there is only a limited number of ways to express an idea, permitting copyright on the expression would effectively monopolize the idea.⁸⁴ In such a case, the expression is said to merge with the idea, and neither is protected by copyright.⁸⁵ In an analogous fashion, seemingly expressive elements dictated by functional demands or efficiency should be excluded from similarity analysis.⁸⁶ The functional end is an idea, and if there are only a few ways to accomplish it, it cannot be protected by copyright.

Erickson v. Blake provides a vivid recent example of the merger doctrine.⁸⁷ There, both the plaintiff and the defendant had the idea to create a musical work based on the digits of π , wherein each digit 0 to 9 was assigned to a note and the decimal expansion of π used to create the theme of a musical work.⁸⁸ *Erickson* held that the usually protectable

Upon any work, and especially upon a play, a great number of patterns of increasing generality will fit equally well, as more and more of the incident is left out. The last may perhaps be no more than the most general statement of what the play is about, and at times might consist only of its title; but there is a point in this series of abstractions where they are no longer protected, since otherwise the playwright could prevent the use of his “ideas,” to which, apart from their expression, his property is never extended. Nobody has ever been able to fix that boundary, and nobody ever can.

Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930) (citations omitted). This formulation, while helpful, leaves unclear where on this spectrum of abstraction copyright ceases to attach.

82. *See* *Boisson v. Banian, Ltd.*, 273 F.3d 262, 272 (2d Cir. 2001).

83. *See* *Shaw v. Lindheim*, 919 F.2d 1353, 1356–57 (9th Cir. 1990).

84. *See* *Concrete Mach. Co. v. Classic Lawn Ornaments, Inc.*, 843 F.2d 600, 606–07 (1st Cir. 1988).

85. *Id.*

86. *See* *Computer Assocs. Int’l, Inc. v. Altai, Inc.*, 982 F.2d 693, 708–10 (2d Cir. 1992).

87. 839 F. Supp. 2d 1132 (D. Or. 2012).

88. *Id.* at 1134.

expression of a melodic sequence had merged with the idea of assigning π 's digits to musical notes, and accordingly found no infringement—all of the similarities between the two works were traceable to this idea.⁸⁹

The related but distinct concept of *scènes à faire*—literally, “scenes which ‘must’ be done”—represents a merger-like doctrine in the artistic context.⁹⁰ *Scènes à faire* are “incidents, characters or settings which are as a practical matter indispensable, or at least standard, in the treatment of a given topic.”⁹¹ For example, a *noir* detective story set in the gritty South Bronx will invariably feature “drunks, prostitutes, vermin and derelict cars.”⁹² These stock elements and fictional conventions merge with general story idea and setting, and are usually—and rightly—excluded from the similarity analysis, regardless of the circuit.⁹³ Outside of the artistic context—particularly in the realm of computer software—courts have extended the notion of *scènes à faire* to include elements dictated by external factors such as compatibility, efficiency, or functionality.⁹⁴

3. *Thin and Thick Protection*

Certain types of works are more likely than others to contain unprotectable elements such as ideas, unoriginal expression, functional elements, stock elements, or elements from the public domain. Courts sometimes speak of these works as having “thin” copyrights.⁹⁵ A thin copyright entails a more exacting standard for similarity—it protects against only “virtually identical” copying.⁹⁶ A simple example is a map. Because much of a map is dictated by its idea—an accurate representation of the territory at issue—maps will typically have a thin copyright vis-à-vis, say, a painting. It will be of little avail in the similarity analysis to observe how both maps depict the same shape of coastline; instead, the plaintiff will have to rely on similarity in expressive choices—selection of what details to include, color combinations, shading—that are not dictated

89. *See id.* at 1139–40.

90. *Schwarz v. Universal Pictures Co.*, 85 F. Supp. 270, 275 (S.D. Cal. 1945).

91. *Hoehling v. Universal City Studios, Inc.*, 618 F.2d 972, 979 (2d Cir. 1980) (quoting *Alexander v. Haley*, 460 F. Supp. 40, 45 (S.D.N.Y. 1978)).

92. *Walker v. Time Life Films, Inc.*, 784 F.2d 44, 50 (2d Cir. 1986).

93. *See* 2 PATRY, *supra* note 75, § 4:24.

94. *See, e.g., Lexmark Int'l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522, 535–36 (6th Cir. 2004). This looser use of *scènes à faire* has been criticized, with some commentators arguing that the doctrine should be limited to the artistic realm. *See, e.g.,* 2 PATRY, *supra* note 75, § 4:26.

95. *See* Balganes, *supra* note 40, at 221–26 (discussing the development of the notion of copyright “thickness”).

96. *Ets-Hokin v. Skyy Spirits Inc.*, 323 F.3d 763, 766 (9th Cir. 2003).

by the facts represented.⁹⁷ While the thin/thick terminology can be helpful, it is really just a kind of shorthand, and it should not be applied mindlessly. Some maps—those commonly found on the inside covers of fantasy novels, for example—are not based upon reality and are thus entitled to a “thicker” copyright.

D. Do the Tests Matter?

The circuits’ various tests for similarity offer plenty of technical jargon for similarity analysis, but what do they mean for plaintiffs and defendants in practice? Are a copyright holder’s chances of success significantly better or worse under, say, the ordinary observer test as opposed to the abstraction/filtration/comparison test? It is not obvious, a priori, which test is more favorable to the copyright holder. On the one hand, it is natural to think that tests applying greater dissection (e.g., the A/F/C and more discerning observer tests) will tend to find less similarity because they make more effort to discount similarities arising from unprotected elements. But on the other hand, a close dissection of the works could reveal *more* copying if expert analysis reveals less-obvious similarities that an ordinary observer might miss.

Thankfully, recent empirical research has begun to shed some light on these long-debated questions. The surprising takeaway of this research is that, for all the attention that the tests have received, they seem to make little difference to case outcomes. When parsed at the finest level, the tests do appear to matter—plaintiff success rates under the A/F/C test are only half those of the more discerning observer test, for example.⁹⁸ But that result is hardly surprising. Since many circuits vary the test based on the type of work at issue (A/F/C for software, discerning observer for thin copyright works),⁹⁹ the comparison is not a fair one. It stands to reason that the difference in outcome might be driven by the different *type of cases* that each test assesses—computer software versus fine arts or musical works—as opposed to the stringency of the test itself. Indeed, the data show huge differences in success rate based on the type of work—ranging from a plaintiff success rate of 53.6% for graphical works to a low of 8.3% for architectural works.¹⁰⁰

97. See *Streetwise Maps, Inc. v. VanDam, Inc.*, 159 F.3d 739, 747–49 (2d Cir. 1998) (noting that “street locations, landmass, bodies of water and landmarks depicted in a map are physical facts” that are unprotected by copyright).

98. See Lippman, *supra* note 2, at 546 fig.5.

99. See *supra* Part I.B.

100. Lippman, *supra* note 2, at 547 fig.6.

A better (though still imperfect) comparison looks to the results under the three main tests: the observer, extrinsic/intrinsic, and A/F/C tests. This comparison ignores the different variants of each test, effectively pooling together the more discerning and ordinary observer, for example. It thus partially removes the confounding factor that many jurisdictions switch between different variants based on the type of work at issue.¹⁰¹ Since several circuits follow each of these major strands of doctrine—the observer test in the First, Second, Third, Fifth, and Seventh Circuits; extrinsic/intrinsic test in the Fourth, Eighth and Ninth; the A/F/C test in the Sixth and Tenth—there is, hopefully, a fair mix of different types of cases analyzed by each test.¹⁰² When this comparison is made, the plaintiff success rates under each of the three main tests are quite similar.¹⁰³ In particular, there is no observed statistical difference between the ordinary observer and extrinsic/intrinsic tests, and plaintiffs are only a few percentage points more likely to succeed under the A/F/C test versus the others.¹⁰⁴ The A/F/C test’s slightly plaintiff-friendly nature may be a result of the fact that computer software cases—which tend to have high success rates—are usually judged under this rubric.¹⁰⁵

II. FRAMING PROBLEMS IN SUBSTANTIAL SIMILARITY

Having reviewed the confused state of substantial similarity doctrine, this Part turns to the primary focus of this Article—the framing problems in substantial similarity. It is motivated by the question left lingering by

101. See *supra* Part I.B.1.

102. See *supra* Part I.B (reviewing the different circuits that follow each test). Another idea would be to compare success rates for each circuit, but this runs into the problem of the dissimilar types of cases heard in each circuit due to the industries under their jurisdiction. The Ninth Circuit, for example, is likely to receive a high number of software cases (because of Silicon Valley) as well as many cases involving Hollywood firms; the Second Circuit likely hears more cases from the world of fine art and publishing. See Lippman, *supra* note 2, at 525. Thus, we should not be surprised to see fairly different success rates across the different circuits, which is in fact what the data reveal. See *id.* at 545 fig.4.

103. See Lippman, *supra* note 2, at 545–46 (“Remarkably, however, when collapsing the six substantial similarity tests into the three main tests—the ordinary observer test, the extrinsic/intrinsic test, and the abstraction-filtration-comparison test—their substantial similarity win rates become closely aligned.”).

104. See Eric Rogers, Comment, *Substantially Unfair: An Empirical Examination of Copyright Substantial Similarity Analysis Among the Federal Circuits*, 2013 MICH. ST. L. REV. 893, 923–24 (footnote omitted) (“[A] plaintiff has the same chances of winning if either the ordinary observer or the extrinsic/intrinsic test regimes are used. However, a plaintiff is statistically less likely to win if a filtration/comparison test is used—5.2% less likely and 3.1% less likely than if the ordinary observer or extrinsic/intrinsic tests are used, respectively.”).

105. See Lippman, *supra* note 2, at 547 fig.6.

the empirical results reviewed above: if the tests for similarity do not make much difference to case outcomes, what does? One possible answer is that the similarity “tests” are less a meaningful analytical tool than a way for judges to justify their ad hoc, intuitive judgment.¹⁰⁶ This may be part of the story, but in a way it only begs the question as to the factors driving similarity analysis. Judges can only indulge in such intuitive decision-making to the extent that the rules for similarity analysis are underspecified.¹⁰⁷ We might reasonably look, then, to aspects of similarity analysis that current doctrine has not settled—those areas in which judges are free to choose the rule that best justifies their decision.

This Part suggests that the ways in which courts frame the comparison between the two works exert a powerful effect on substantial similarity outcomes. Evidence from the cognitive sciences reveals that framing effects can impact perceived similarity. One of the core findings of this literature is that the perspective of the comparison is critical: observers are found to perceive similarity differently depending on which of the two items compared is taken as the subject, and which as the baseline. In particular, people perceive greater similarity when the baseline is a more prominent concept; they are more likely to agree that “an ellipse is like a circle” than “a circle is like an ellipse.”¹⁰⁸ As Amos Tversky put it succinctly: “[w]e say ‘the son resembles the father’ rather than ‘the father resembles the son.’”¹⁰⁹

Issues of perspective and framing are therefore of potentially great importance to copyright’s similarity analysis. Using several examples from the case law, this Part illustrates how courts can exploit the lack of consensus on these issues to frame the inquiry in ways that tilt the comparison in favor of the plaintiff or defendant. The first Part examines similarity’s timing problem, the question of when the idea-expression distinction should apply. The second Part looks at how courts choose the appropriate baseline for similarity, using either the original work or the accused work (or some other measure) as the benchmark. The final Part turns to the question of whether notable dissimilarities are considered in the comparison.

106. See Said, *supra* note 12, at 521–23 (criticizing purely intuitive methods of interpretation in copyright as manipulable and lacking transparency); Stanfield, *supra* note 12, at 492, 512.

107. Cf. Said, *supra* note 12, at 515–23 (arguing that copyright doctrine should do more to restrain judges’ discretion in interpretative choices, including infringement analysis).

108. Tversky, *supra* note 13, at 328, 333–34.

109. *Id.* at 328.

A few key points emerge from the analysis. The first is framing's relationship with outcomes. How a court chooses to frame the comparison is seen to be strongly associated with whether substantial similarity is ultimately found. It is important not to infer too much about causality, however. It may be that the framing choices bias the judicial comparison in favor of or against similarity. Or, perhaps more plausibly, courts might select the particular framing that puts their desired outcome in the best light. The second observation is that framing problems often relate to the question of perspective. Whether we are discussing issues of timing, baseline, or dissimilarities, one core issue looms: Should the factfinder assume the viewpoint of the alleged infringer, or of the original creator?

A. *The Timing Problem*

This Part examines several instances of similarity's timing problem. At bottom, the timing problem represents confusion about how to apply the idea-expression distinction in a dynamic context. Expression that was original and protected by copyright when created may subsequently become (or merge with) an "idea, procedure, process, system, method of operation, concept, principle, or discovery" that would not ordinarily be protected by copyright.¹¹⁰ The question, then, is whether to apply the idea-expression distinction at the time of the original work's creation, or at the time of the alleged infringement.

The timing problem represents a particular concern for material that becomes a de facto or industry standard after its creation.¹¹¹ For example, a model building code is initially a copyrightable literary work, but once a jurisdiction adopts it as law, it effectively loses its protection: citizens need to be able to state what the law is, and there is no other way to express that idea but to quote from the code.¹¹² In the context of computer software, a split has developed over a related issue: if a piece of code or a

110. 17 U.S.C. § 102(b) (2014); *see also supra* Part I.C.1–2 (discussing the idea-expression distinction and the merger doctrine).

111. *See generally* Pamela Samuelson, *Questioning Copyright in Standards*, 48 B.C. L. REV. 193 (2007) (presenting examples of this problem and arguing that standards should not be protected by copyright); Stephen Preonas, *Mergicide, When Good Copyrights Go Bad: A Recommendation for a Market-Based, Defendant-Centric Approach to the Merger Doctrine in the Context of Compilations*, 11 INTELL. PROP. L. BULL. 89, 97–109 (2006) (proposing loss of copyright for elements that become standards); *see also* Peter S. Menell, *An Analysis of the Scope of Copyright Protection for Application Programs*, 41 STAN. L. REV. 1045, 1098–1102 (1989) (arguing for limiting copyright de facto standard elements in software, including user interfaces and elements necessary for compatibility); Timothy S. Teter, Note, *Merger and the Machines: An Analysis of the Pro-Compatibility Trend in Computer Software Copyright Cases*, 45 STAN. L. REV. 1061, 1066–72 (1993) (same).

112. *See* *Veeck v. S. Bldg. Code Cong. Int'l, Inc.*, 293 F.3d 791, 801–02 (5th Cir. 2002).

user interface becomes so widespread as to become a functional standard, should it still be protected by copyright?¹¹³ Sometimes, courts discuss this issue as whether the doctrines of merger and *scènes à faire* should be assessed as part of copyrightability, or as part of the infringement inquiry.¹¹⁴ Either way, it amounts to the same problem: Should we look to the choices presented to the original creator, and protect the material so long as it was copyrightable at the time of its creation? Or should we instead assume the viewpoint of the defendant, and look to the choices available to her when the allegedly infringing work was created?

Oracle America, Inc. v. Google Inc., the most important software copyright decision in many years, presents a clear illustration of the different perspectives on the timing problem.¹¹⁵ Both the district and appellate courts in *Oracle* applied Ninth Circuit precedent, and thus both decisions were governed by the same legal framework for substantial similarity: the A/F/C test.¹¹⁶ However, their distinct approaches to the timing problem—particularly, whether one assumes the perspective of the original creator or alleged infringer—led the courts to starkly different conclusions.

The dispute in *Oracle* requires a bit of technical background to appreciate. In creating its wildly successful Android operating system for mobile phones, Google copied certain aspects of Oracle's application programming interfaces ("APIs") for the Java computer programming language.¹¹⁷ APIs are collections of "shortcuts" within Java that programmers can call upon to perform common functions, instead of writing basic code from scratch.¹¹⁸ For example, one API contains a class called "math," within which there is a "max" method that programmers

113. *Compare, e.g., Lotus Dev. Corp. v. Borland Int'l, Inc.*, 49 F.3d 807, 816 (1st Cir. 1995) (holding that spreadsheet menu hierarchy that became common user standard lacked copyright protection as a "method of operation"), *aff'd by equally divided court*, 516 U.S. 233 (1996) (per curiam), *with Mitel, Inc. v. Iqtel, Inc.*, 124 F.3d 1366, 1372 (10th Cir. 1997) (concluding that telecommunications standard command codes contained protectable expression despite being a "method of operation"), *and Oracle Am., Inc. v. Google Inc.*, 750 F.3d 1339, 1372 (Fed. Cir. 2014) (rejecting proposition that elements that become an "effective industry standard" lose copyright protection).

114. *See Lexmark Int'l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522, 557–59 (6th Cir. 2004) (Feikens, J., concurring in part and dissenting in part) (reviewing split in authority as to whether merger and *scènes à faire* doctrines are assessed at the copyrightability stage or the infringement stage).

115. 750 F.3d 1339 (Fed. Cir. 2014), *cert. denied*, 135 S. Ct. 2887 (2015).

116. *Id.* at 1357–58. Because this district court viewed the issue as primarily one of copyrightability, it did not reach the infringement issue directly, though it discussed and implied the principles of the filtration step. *See id.*

117. *Id.* at 1347.

118. *Id.* at 1348–49.

can use to output the greater of two input values.¹¹⁹ In building Android, Google copied the “declaring code” of 37 out of the 166 API packages in Java.¹²⁰ The declaring code is simply the name for the function (in our example, “max”) and its definition, as well as its place within the APIs’ organization (in our example, within the “math” class).¹²¹ Google independently wrote implementing code—the operative aspects—for each API.¹²²

In short, Google copied the “names” and definitions of thirty-seven API packages, as well as aspects of the overall organization of the APIs. Google claimed it was entitled to do this in order to allow programmers to make functional use of Java in creating apps for Android—and the parties acknowledged that the Java language itself was open for all to use.¹²³ In other words, Google claimed that it copied the APIs not because of their elegant expression—that “max” was a particularly apt name—but because that terminology had become a standard with which Java programmers were familiar. The district court agreed and ruled in favor of Google, concluding that the declaring code was not copyrightable, due (in significant part) to these interoperability concerns.¹²⁴ The Federal Circuit emphatically and unanimously reversed.¹²⁵

The competing approaches of the district and appeals courts in *Oracle* vividly illustrate the lack of judicial consensus as to when to apply the idea-expression distinction, and its stark effect. The district court took the view that the line between uncopyrightable idea and protectable expression should be determined at the time of infringement, and looked to the choices available to *Google* when it copied the code. It acknowledged that there may have been creative choices made in the original design of the Java APIs.¹²⁶ But, once designed and widely used, anyone who wished to make practical use of the Java language had no choice other than to use the names and structure that were standard and familiar to programmers.¹²⁷

119. *Id.* at 1349–50.

120. *Id.* at 1348–49.

121. *Id.* at 1350–51.

122. *Id.*

123. See *Oracle Am., Inc. v. Google Inc. (Oracle District Court)*, 872 F. Supp. 2d 974, 978 (N.D. Cal. 2012) (“All agree that Google was and remains free to use the Java language itself.”), *rev’d*, 750 F.3d 1339 (Fed. Cir. 2014).

124. *Id.* at 998–1002.

125. *Oracle*, 750 F.3d at 1381.

126. See *Oracle District Court*, 872 F. Supp. 2d at 997–98.

127. See *id.* at 1000 (emphasis omitted) (“In order for at least some of this code to run on Android, Google was required to provide the same `java.package.Class.method()` command system using the

The Federal Circuit held that this was error. It reasoned that the declaring code was protected by copyright because the names involved some creative choice at the time of Java's creation:

We further find that the district court erred in focusing its merger analysis on the options available to Google at the time of copying. It is well-established that copyrightability and the scope of protectable activity are to be evaluated at the time of creation The focus is, therefore, on the options that were available to Sun/Oracle at the time it created the API packages.¹²⁸

In other words, because the function called "max" could equally have been called "greatest" or "most" when it was created, it was original and protected by copyright. The fact that "max" subsequently became a widespread standard was irrelevant to the infringement analysis.¹²⁹ As these subsequent developments were at the core of Google's justification for copying, the Federal Circuit's approach to the timing problem essentially determined the outcome in *Oracle*.

The Sixth Circuit took a quite different approach to similarity's timing problem in *Lexmark International, Inc. v. Static Control Components, Inc.*¹³⁰ In fact, the *Lexmark* case is in itself a fascinating study of the issue, offering three distinct approaches to the timing issue as embodied in the rationales of the district court, the majority appellate opinion by Judge Sutton, and the concurrence of Judge Feikens.

Lexmark is a leading manufacturer of laser and inkjet printers.¹³¹ Like most printer manufacturers, it also sells toner cartridges for its printers; in an effort to boost its control over the cartridge market, *Lexmark* sells what it calls "prebate" cartridges.¹³² In return for a reduced price, the consumer agrees not to refill and reuse the toner cartridge via a third party, but instead to return the cartridge to *Lexmark*.¹³³ To enforce this, *Lexmark* includes a microchip on its prebate cartridges that uses computer code—

same names with the same 'taxonomy' and with the same functional specifications. Google replicated what was necessary to achieve a degree of interoperability—but no more, taking care, as said before, to provide its own implementations.").

128. *Oracle*, 750 F.3d at 1361 (citation omitted).

129. *See id.* at 1372 ("[T]o the extent Google suggests that it was entitled to copy the Java API packages because they had become the effective industry standard, we are unpersuaded. Google cites no authority for its suggestion that copyrighted works lose protection when they become popular, and we have found none.").

130. 387 F.3d 522 (6th Cir. 2004).

131. *Id.* at 529.

132. *Id.* at 530.

133. *Id.*

Lexmark's "Toner Loading Program"—to disable the cartridge once it runs out of toner.¹³⁴ Static Control Components ("SCC") engineered a microchip that mimics Lexmark's code to allow consumers to refill and reuse their cartridges; in effect, SCC's chip copies Lexmark's code in order to enable third-party manufacturers to make and sell refurbished toner cartridges.¹³⁵ Lexmark was not pleased about this and sued SCC for copyright infringement and violations of the Digital Millennium Copyright Act.¹³⁶

One way to think of Lexmark's Toner Loading Program ("TLP") is as a kind of "self-creating" standard. Stripped of context, the TLP is just a simple piece of code that estimates the toner level.¹³⁷ But once Lexmark uses the code in the particular way it did—the TLP is an input in the authentication sequence necessary to make the printer function¹³⁸—Lexmark has transformed the TLP into a functional standard or "method of operation" outside the scope of copyright.¹³⁹ The district court viewed matters differently, however, and it issued a preliminary injunction against SCC.¹⁴⁰ In essence, the district court took the same approach to the timing problem as the Federal Circuit did in *Oracle*. It assessed the idea-expression distinction at the time of the TLP's creation, and simply asked whether it was possible to write different code that would perform the same function as the TLP.¹⁴¹ Finding that a toner loading program might be written "in a number of different ways," the district court ruled that the TLP was expression entitled to copyright protection, which SCC duly infringed.¹⁴²

The Sixth Circuit reversed, finding that Lexmark was not likely to succeed on its claim of copyright infringement.¹⁴³ Writing for the majority, Judge Sutton focused not on infringement, but on originality—whether the TLP should have any copyright at all. As to the timing problem, Judge Sutton looked at the issue from SCC's perspective, emphasizing how

134. *Id.*

135. *Id.* at 530–31.

136. *Id.* at 531.

137. *See id.* at 529–30.

138. *See id.* at 541.

139. *See* 17 U.S.C. § 102(b) (2014).

140. *Lexmark Int'l, Inc. v. Static Control Components, Inc. (Lexmark District Court)*, 253 F. Supp. 2d 943, 974 (E.D. Ky. 2003), *vacated*, 387 F.3d 522 (6th Cir. 2004).

141. *See id.* at 950–51 ("The Toner Loading Programs may be expressed in different ways to perform the same function, namely estimating the amount of toner . . . [T]he Toner Loading Programs [thus] constitute[] creative expression and [are] entitled to copyright protection.").

142. *Id.* at 962–65.

143. *Lexmark*, 387 F.3d at 529.

much of the TLP, as used, was dictated by efficiency and functional constraints.¹⁴⁴ Applying the doctrine of merger and *scènes à faire* in light of these constraints, Judge Sutton suggested that the TLP lacked sufficient originality to be copyrightable.¹⁴⁵

Judge Feikens, concurring in part and dissenting in part, suggested that the TLP may be copyrightable, but that SCC did not infringe.¹⁴⁶ Judge Feikens first noted a split in the circuits as to whether the doctrines of merger and *scènes à faire* act as a bar to copyrightability, or instead as a part of the substantial similarity inquiry.¹⁴⁷ The practical difference between these two routes is that the former would remove all copyright protection for the work, whereas the latter simply states that the work is not protected as used in a particular context. Judge Feikens argued that the latter was the sounder course, making a helpful analogy:

[A]n otherwise copyrightable text can be used as a method of operation of a computer—for instance, an original, copyrightable poem could be used as a password, or a computer program as a lock-out code. In my view, therefore, it is necessary to know what the potential infringer is doing with the material in order to know if merger has occurred. . . .

Under this reasoning, an individual who copied a poem solely to use as a password would not have infringed the copyright, because in that scenario, the alleged infringer would have the defense that the poem has “merged” with a method of operation (the password). By contrast, someone who copied the poem for expressive purposes (for instance, as part of a book of poetry) would not have this defense.¹⁴⁸

Judge Feikens then turned to the infringement analysis. Like Judge Sutton, he assessed the idea-expression distinction at the time of infringement, looking not to the choices of Lexmark when it wrote the TLP, but those of SCC when it copied the TLP. Judge Feikens suggested that, so long as

144. *See id.* at 537 (holding that the district court erred in “refusing to consider whether ‘external factors such as compatibility requirements, industry standards, and efficiency’ circumscribed the number of forms that the Toner Loading Program could take”).

145. *Id.* at 541.

146. *Id.* at 558 (Feikens, J., concurring in part and dissenting in part).

147. *See id.* at 557, 559 (explaining and citing cases on both sides of these circuit splits).

148. *Id.* at 557–58.

SCC used the TLP in a functional manner—to interface with Lexmark’s printers—it would not infringe.¹⁴⁹

The sharp contrast between the approaches in *Lexmark* and *Oracle* illustrates the unsettled nature of the timing problem and its impact on case outcomes. Despite the Federal Circuit’s blithe assertion that the timing issue is “well-settled,” actual examination of the case law reveals decisions going in both directions.¹⁵⁰ Especially in the context of software, where material can quickly become a widely adopted standard, when a court chooses to assess the idea-expression distinction has a huge impact on whether similarity is ultimately found.¹⁵¹

B. The Baseline Problem

This Part explores similarity’s baseline problem: how courts determine the benchmark used to measure when similarity becomes “substantial,” and therefore an infringement of copyright. In assessing the quantitative significance of what the alleged infringer copied, should we look to how much was copied from the original work, or how much of the accused work is copied material? Or should we instead assess the significance against some absolute threshold, or solely as a qualitative matter—the economic, artistic, or practical significance of the material copied? In contrast to the timing problem, where there are two main camps—evaluating idea-expression at the time of creation, versus the time of infringement—courts take a looser approach to the baseline for infringement. To the extent there is any consensus, the traditional rule is that the appropriate baseline is the original work.¹⁵² However, courts’

149. *See id.* at 558 (“Defendant can still avoid infringement, however, if it uses the TLP only as a method of operation. For instance, Defendant can only claim this defense to infringement if it uses the TLP to interface with the Lexmark printers at issue, and if it is a necessary method of operation of the machine.”).

150. In addition to the contrast between *Lexmark* and *Oracle*, compare, e.g., *Matthew Bender & Co. v. W. Publ’g Co.*, 158 F.3d 693, 701 (2d Cir. 1998) (holding that standard pagination in Westlaw case reporters was not protected by copyright), with *W. Publ’g Co. v. Mead Data Cent., Inc.*, 799 F.2d 1219, 1228 (8th Cir. 1986) (holding that standard pagination in Westlaw case reporters was protected by copyright). For further discussion, see generally Sandro Ocasio, Comment, *Pruning Paracopyright Protections: Why Courts Should Apply the Merger and Scènes à Faire Doctrines at the Copyrightability Stage of the Copyright Infringement Analysis*, 3 SETON HALL CIRCUIT REV. 303, 310–23 (reviewing circuit split on the issue of whether merger and *scènes à faire* are assessed as part of copyrightability or infringement).

151. *See generally* Menell, *supra* note 111, at 1066–67 (arguing that network effects are particularly salient in the software context, often leading to standardization); Teter, *supra* note 111, at 1066–70 (arguing that standardization can create “lock-in” effects for users and that copyright accordingly should not protect elements required for compatibility or de facto standards).

152. *See, e.g.*, *Sheldon v. Metro-Goldwyn Pictures Corp.*, 81 F.2d 49, 56 (2d Cir. 1936).

choice of baseline is often implicit, and courts frequently employ whatever baseline provides greater rhetorical support for their argument.

Although the case arose in the context of fair use, the Supreme Court's decision in *Harper & Row, Publishers, Inc. v. Nation Enterprises* provides a simple example to illustrate the baseline problem.¹⁵³ *Harper & Row* addressed whether The Nation magazine's "scoop" of President Ford's unpublished memoir, *A Time to Heal*, was a fair use.¹⁵⁴ Having received a prepublication copy from an unidentified source, The Nation quoted and paraphrased from the forthcoming memoir to produce an article detailing Ford's decision to pardon former President Richard Nixon.¹⁵⁵ The Nation's article was 2250 words (about five to ten pages), of which about 300 were direct quotes from the Ford memoir.¹⁵⁶ The original work, Ford's memoir, was about 200,000 words.¹⁵⁷

The third statutory factor for determining whether a use is fair directs courts to consider "the amount and substantiality of the portion used in relation to the copyrighted work as a whole."¹⁵⁸ This suggests that the proper baseline in the fair use context is the original work, and the majority opinion in *Harper & Row* pays some lip service to that notion.¹⁵⁹ Taken literally, this measure would suggest that what The Nation took was quantitatively trivial—300 words out of 200,000, or about 0.1%. Justice O'Connor's majority opinion concedes that 300 words is an "insubstantial" part of Ford's memoir.¹⁶⁰ Rather abruptly, however, the majority opinion shifts the baseline and emphasizes that the quotes were "at least 13% of [The Nation's] infringing article"—and, moreover, very significant as a qualitative matter: the "heart of the book."¹⁶¹ Justice Brennan's dissent, in a footnote, gently takes the majority to task for this switch.¹⁶²

153. 471 U.S. 539 (1985).

154. *Id.* at 542–43.

155. *Id.*

156. *Id.* at 545.

157. *Id.* at 579 (Brennan, J., dissenting).

158. 17 U.S.C. § 107(3) (2014).

159. *See Harper & Row*, 471 U.S. at 564 ("[T]he Act directs us to examine the amount and substantiality of the portion used in relation to the copyrighted work as a whole.").

160. *Id.*

161. *Id.* at 564–65.

162. *Id.* at 599–600 n.23 (Brennan, J., dissenting) ("Superficially, the Court would thus appear to be evaluating The Nation's quotation of 300 words in relation to the amount and substantiality of expression used in relation to the second author's work as a whole. The statute directs the inquiry into 'the amount and substantiality of the portion used in relation to *the copyrighted work* as a whole,' 17 U.S.C. § 107(3) (emphasis added). As the statutory directive implies, it matters little whether the second author's use is 1- or 100-percent appropriated expression . . .").

Substantial similarity, even more so than fair use, is a common law doctrine, and its baseline is not dictated by any statutory considerations. In practice, however, courts tend to follow the course of the Supreme Court in *Harper & Row*. They pay lip service to the notion that the original work is the appropriate baseline—if they address the issue at all—but then disregard analytical consistency to select whatever baseline strengthens their argument.

Newton v. Diamond provides an atypically thorough examination of the baseline problem in the substantial similarity context.¹⁶³ *Newton* involved the Beastie Boys' song "Pass the Mic," which prominently featured a sample of "Choir" by the jazz flutist James Newton.¹⁶⁴ Newton's technique in "Choir" was unusual, with the score instructing the performer to "sing[] into" the flute while fingering particular pitches, creating a ghostly effect.¹⁶⁵ The Beastie Boys took a six-second sample from "Choir" and looped it, using the sample as a background element for almost all of the duration of "Pass the Mic."¹⁶⁶ Newton sued for copyright infringement, and the district court granted summary judgment in favor of the Beastie Boys.¹⁶⁷

The case is complicated somewhat by the unusual nature of Newton's work, and the fact that the Beastie Boys had a partial license to use it. There are two types of copyrights in a musical recording: one for the composition, and another for the sound recording itself.¹⁶⁸ The former covers the work of the writer of the music and lyrics, which is sometimes embodied in a score; the latter covers the work of the performers and producers of a particular recording.¹⁶⁹ For "Pass the Mic," the Beastie Boys secured permission from the owner of the sound recording copyright (which Newton had assigned), but not from the owner of the composition copyright, which remained with Newton.¹⁷⁰ Thus, some of the debate in the case focused on delineating which aspects of the sample were attributable to Newton's composition, and which to his performance.¹⁷¹

163. 388 F.3d 1189 (9th Cir. 2003).

164. *Id.* at 1191.

165. *Id.* at 1197–99 (Graber, J., dissenting).

166. *Id.* at 1192.

167. *Id.*

168. 17 U.S.C. § 102(a)(2), (7) (2014).

169. *See* Griffin v. J-Records, 398 F. Supp. 2d 1137, 1142–43 (E.D. Wash. 2005) (discussing distinctions between copyright in musical compositions and sound recordings); *see also generally* U.S. COPYRIGHT OFFICE, CIRCULAR 56A: COPYRIGHT REGISTRATION OF MUSICAL COMPOSITIONS AND SOUND RECORDINGS 1 (2012), *available at* <http://copyright.gov/circs/circ56a.pdf> (same).

170. *Newton*, 388 F.3d at 1191.

171. *Compare id.* at 1195 (characterizing the sample as a "three-note sequence"), *with id.* at 1197–

Putting that issue aside, the judges in *Newton* struggled to define the baseline against which to measure whether the Beastie Boys' use of the sample was substantial enough to be an infringement.¹⁷² On the one hand, the sample was a limited part of Newton's work (only a few seconds), but on the other, the sample was a significant part of the infringing work (the main background element). The majority in *Newton* cited the traditional rule that the plaintiff's work should be the baseline, and suggested that the copying was insubstantial as a result.¹⁷³ It emphasized that the sample "appears only once in Newton's composition," and that six seconds was merely "two percent of the four-and-a-half-minute" running time of "Choir."¹⁷⁴ But what of the fact that the Beastie Boys looped the sample, such that it constituted a large part of the accused work? Strikingly, because of its chosen baseline—and in contrast to the approach in *Harper & Row*—the majority found that the looping was entirely "irrelevant in weighing the sample's qualitative and quantitative significance."¹⁷⁵

It is worthwhile to trace how *Newton* reached its conclusion to disregard how the defendants actually used the sample when assessing similarity. The rationale for assessing infringement against the baseline of the original work derives—as does a surprising amount of copyright doctrine—from an opinion by Judge Learned Hand. In *Sheldon v. Metro-Goldwyn Pictures, Corp.*, Judge Hand addressed whether a play and a movie—both loosely based on the same real-life story of Madeleine Smith, a cause célèbre of the day—were similar enough in plot, character, and incident such that the movie infringed the earlier play.¹⁷⁶ Although the movie took no dialogue directly from the play, Judge Hand thought the characters and certain scenes were so alike that the movie was an infringement.¹⁷⁷ Rejecting out of hand the defendant's attempt to rely on the many differences between the movie and picture, Judge Hand held that this was "immaterial" because "no plagiarist can excuse the wrong by showing how much of his work he did not pirate."¹⁷⁸

98 (Graber, J., dissenting) (observing that the sequence is sung above the fingered notes, creating "four separate tones," and that the unusual playing technique was transcribed in the score). The majority conceded, for purposes of the appeal, that the "multiphonics" obtained through Newton's technique were part of the composition. *Id.* at 1191.

172. *Id.* at 1192–96.

173. *Id.* at 1195 (citing *Sheldon v. Metro-Goldwyn Pictures Corp.*, 81 F.2d 49, 56 (2d Cir. 1936)).

174. *Id.* at 1195–96.

175. *Id.* at 1195 (emphasis added).

176. 81 F.2d 49, 49–53.

177. *Id.* at 54–55.

178. *Id.* at 56.

Judge Hand’s pithy line—which I will call “the *Sheldon* rule”—has become something of a mantra in copyright infringement cases, being “wildly over-cited” in subsequent decisions.¹⁷⁹ Interestingly, the *Sheldon* rule was originally designed to benefit the copyright holder by ignoring any new material that the infringer introduced. The concern motivating Judge Hand seems to be that an infringer could escape liability simply by “tacking on” a large amount of unrelated new material. The majority in *Newton*, however, used *Sheldon* in quite the opposite way, for the benefit of the defendant. It discounted how the Beastie Boys actually used the sample, focusing on the fact that the sample was merely a small part of the original work. Dissenting in *Newton*, Judge Graber felt that the majority had unfairly dismissed the qualitative importance of what was copied from *Newton*, noting that much of what made *Newton*’s work so distinctive was the playing technique, a part of the composition and of the sample.¹⁸⁰

The recent case of *Antonick v. Electronic Arts Inc.*¹⁸¹ suggests that a literal application of the *Sheldon* rule can sometimes produce troubling results. Robin Antonick, pursuant to an agreement with Electronic Arts (“EA”), designed and wrote source code for 1988’s *John Madden Football*, a video game for the Apple II computer.¹⁸² That game was the first in a phenomenally successful franchise; later games for systems like the Sega Genesis, Xbox, and PlayStation have sold over 85 million copies and grossed around \$4 billion for EA.¹⁸³ Antonick’s contractual agreement with EA entitled him to royalties on any “derivative works” using his code, and—after discovering similarities between the first game and later entries in the *Madden* series—Antonick sued EA for royalties based on the sale of the later games.¹⁸⁴ EA claimed that the code for the later games was developed independently—it had, falsely, told Antonick the same thing for years.¹⁸⁵ The jury, however, found that EA copied Antonick’s code for the

179. See, e.g., 3 PATRY, *supra* note 5, § 9:65. A Westlaw search for the quote yields ninety-five federal cases citing Hand’s statement verbatim.

180. See *Newton*, 388 F.3d at 1197–98 (Graber, J., dissenting). Because she considered the sample to be qualitatively significant, Judge Graber declined to assess its quantitative significance. *Id.* at 1198 n.3.

181. No. C 11-1543 CRB, 2014 WL 245018 (N.D. Cal. Jan. 22, 2014).

182. *Id.* at *1.

183. See, e.g., Joel Rosenblatt, *Electronic Arts Must Pay \$11 Million After Verdict*, BLOOMBERG (July 23, 2013, 8:52 PM), <http://www.bloomberg.com/news/articles/2013-07-24/electronic-arts-must-pay-11-million-after-verdict>.

184. *Antonick v. Electronic Arts Inc.*, No. C 11-01543 CRB, 2011 WL 4501324, at *2–3 (N.D. Cal. Sept. 27, 2011).

185. Indeed, the long delay in bringing suit was, according to Antonick, the result of misrepresentations by EA that the later versions of *Madden Football* were developed in a “clean

football plays and formations and used this code in the later games; it awarded Antonick \$11 million in damages.¹⁸⁶

EA sought a new trial, arguing that the first *Madden* game was not substantially similar to the later versions as a matter of law—and therefore no royalties were owed.¹⁸⁷ Judge Breyer agreed that the jury was entitled to find that EA's code for the later *Madden* games governing plays and formations was, in fact, copied from Antonick.¹⁸⁸ But he ruled that this copying was not substantial when comparing the original and accused games “as a whole.”¹⁸⁹ The code that EA had copied from Antonick was only one part of a larger video game. The jury was required to find that the later versions were similar enough on the level of the entire work—not just in the infringing module.¹⁹⁰ This, Judge Breyer claimed, no reasonable jury could find based on the evidence presented.¹⁹¹ Confusingly, although Judge Breyer cited *Newton* to say that significance is to be measured against the plaintiff's work, he primarily faulted Antonick for not presenting evidence of the whole of EA's later *Madden* games (the accused works).¹⁹²

The result in *Antonick* illustrates some potential problems with a strict application of *Sheldon*'s plaintiff's-work-as-baseline rule. Why should EA be excused from liability just because the code it copied from Antonick happened to be incorporated into a large work? Would EA's copying of the plays and formations code be any more or less wrongful if the *Madden* games had been “smaller” works—say, a text-based computer game? If, as *Sheldon* suggests, the fact that the accused work contains extra expression should not negate substantial copying, then why should the original work not get the same benefit? It seems odd that Antonick should be penalized, in effect, for creating a more complex video game.

room” without any use of his code. *Id.* at *1. The jury credited Antonick's version of events. *See Antonick*, 2014 WL 245018, at *5.

186. Rosenblatt, *supra* note 183.

187. Antonick's contract with EA defined the term “derivative works” to parallel the statutory definition in US copyright law; thus, the contractual question is identical to the usual legal one of whether the allegedly infringing works were “substantially similar” to the original Apple II game. *See Antonick*, 2014 WL 245018, at *6.

188. *Id.*

189. *Id.* at *7–10.

190. *Id.* at *6–7.

191. *Id.* at *7–10. Antonick has appealed the decision and the case is pending in the Ninth Circuit. *See Antonick v. Electronic Arts Inc.*, No. C 11-1543 CRB, 2014 WL 245018 (N.D. Cal. Jan. 22, 2014), *appeal docketed*, No. 14-15298 (9th Cir. Feb. 19, 2014).

192. *See Antonick*, 2014 WL 245018, at *7 (“The Jury Had No Evidence of Sega Madden as a Whole.”).

The *Oracle v. Google* dispute, discussed above,¹⁹³ illustrates yet another approach to substantial similarity's baseline, and one at odds with both *Newton* and *Antonick*. In addition to their dispute over the timing problem, the district and appellate courts in *Oracle* also differed on how to assess the substantiality of what Google copied. The district court followed *Sheldon* and emphasized that very little code was copied in the context of the original work: of 166 Java APIs, 129 were not copied at all, and of the 37 at issue, only 3% of the API code (the declaring code) was copied.¹⁹⁴ In fact, 3% was a generous estimate: as both Android and the Java platform writ large contain millions of lines of code, the amount copied is quantitatively infinitesimal regardless of which work is used as the measure.¹⁹⁵ The Federal Circuit in *Oracle*, however, did not let that stand in its way. Disregarding both baselines, it emphasized the total amount copied as an *absolute* matter, repeatedly referring to the “7,000 lines of code” that Google copied.¹⁹⁶

In sum, courts have adopted a shifting approach to similarity's baseline—sometimes paying lip service to the *Sheldon* rule, but deviating from it when it is helpful to make a rhetorical point, strengthen an argument, or avoid a thorny issue. The result is unpredictability in how to assess quantitative significance in copyright infringement.

C. The Relevance of Dissimilarity

This Part analyzes several examples of how dissimilarities are considered—or not—in substantial similarity analysis. On the surface, the doctrine in this area appears to be at odds with itself. Citing the *Sheldon* rule, some courts refuse to consider dissimilarities, lest the defendant escape liability “by showing how much of [the] work he did not pirate.”¹⁹⁷ The influential Nimmer copyright treatise takes *Sheldon* to an extreme, asserting that “[i]t is entirely immaterial that, in many respects, plaintiff's and defendant's works are dissimilar, if in other respects, similarity as to a

193. See *supra* notes 115–29 and accompanying text.

194. *Oracle Am., Inc. v. Google Inc. (Oracle District Court)*, 872 F. Supp. 2d 974, 1001 (N.D. Cal. 2012) (“In closing, it is important to step back and take in the breadth of Oracle's claim. Of the 166 Java packages, 129 were not violated in any way. Of the 37 accused, 97 percent of the Android lines were new from Google and the remaining three percent were freely replicable under the merger and names doctrines.”), *rev'd*, 750 F.3d 1339 (Fed. Cir. 2014).

195. See *Oracle Am., Inc. v. Google Inc.* 750 F.3d 1339, 1379 (Fed. Cir. 2014).

196. See *id.* at 1356, 1359, 1363.

197. See, e.g., *Tufenkian Imp./Exp. Ventures, Inc. v. Einstein Moomjy, Inc.*, 338 F.3d 127, 132 (2d Cir. 2003) (quoting *Sheldon v. Metro-Goldwyn Pictures Corp.*, 81 F.2d 49, 56 (2d Cir. 1936)).

substantial element of plaintiff's work can be shown."¹⁹⁸ But courts just as easily invoke another widely cited maxim: "[n]umerous differences tend to undercut substantial similarity."¹⁹⁹ The tension between these two lines of jurisprudence creates uncertainty about the significance of dissimilarities in infringement analysis.

A careful treatment of this issue can be found in *Warner Bros. v. American Broadcasting Companies*.²⁰⁰ That case presented questions about the scope of the copyright in the character of Superman. ABC's television show *The Greatest American Hero* featured a character, Ralph Hinkley, who has a number of superficial features in common with Superman.²⁰¹ Hinkley wears a skin-tight leotard with a prominent insignia on his chest, and a cape; he has superhuman speed and strength derived from an extraterrestrial source; he can fly; he is impervious to bullets; he has "holographic" (though not x-ray) vision.²⁰² Moreover, characters on the show knowingly reference the Superman franchise. Looking at himself in the mirror, Hinkley says, "It's a bird . . . it's a plane . . . it's Ralph Hinkley."²⁰³ After watching his unsuccessful attempt at flight, a youngster tells Hinkley, "Superman wouldn't do it that way."²⁰⁴

If Nimmer's view of the *Sheldon* rule is taken at face value, we might think that these many similarities would be the end of the similarity analysis. But Judge Newman's opinion in *Warner Bros.* found that Hinkley's *differences* with Superman were important, and undermined the similarities. Hinkley is a bumbling, awkward superhero—an inversion of the straitlaced, confident Superman. Hinkley lost the instruction manual that came with his superpowers; when he tries to fly, he cannot steer and crash lands; he is impervious to bullets, but still cowers and covers his face when shot at; his suit is red, not blue.²⁰⁵ Are we to disregard these dissimilarities?

In *Warner Bros.*, Judge Newman first observed the apparent "paradox" between the *Sheldon* rule and the notion that a defendant may avoid infringement by making sufficient changes to her work:

198. NIMMER, *supra* note 24, § 13.03[B][1][a].

199. *Durham Indus., Inc. v. Tomy Corp.*, 630 F.2d 905, 913 (2d Cir. 1980) (citing *Herbert Rosenthal Jewelry Corp. v. Honora Jewelry Co.*, 509 F.2d 64, 65 (2d Cir. 1974)).

200. 720 F.2d 231 (2d Cir. 1983).

201. *Id.* at 236.

202. *Id.* at 236–37.

203. *Id.* at 237.

204. *Id.*

205. *Id.* at 236–38.

The two propositions are not facially inconsistent; the second proposition contemplates a work that *would be* substantially similar if its author had not made changes from the plaintiff's work. Yet in practice the distinction between the two propositions has become somewhat blurred. We have observed that "numerous differences tend to undercut substantial similarity." . . .

The tension between these two propositions perhaps results from their formulation in the context of literary works and their subsequent application to graphic and three-dimensional works. A story has a linear dimension: it begins, continues, and ends. If a defendant copies substantial portions of a plaintiff's sequence of events, he does not escape infringement by adding original episodes somewhere along the line. A graphic or three-dimensional work is created to be perceived as an entirety. Significant dissimilarities between two works of this sort inevitably lessen the similarity that would otherwise exist²⁰⁶

In other words, the *Sheldon* rule only makes sense, if at all, when the defendant has tacked on *additional* expression to an already-infringing work. It should not be applied blindly to disregard important differences between two works, and especially not to "non-linear" works such as characters or graphical works. Finding that the striking differences between Ralph Hinkley and Superman in personality undercut their similarity in superpowers, Judge Newman found no substantial similarity.²⁰⁷

Judge Newman is certainly correct that the *Sheldon* rule can be understood narrowly to permit consideration of dissimilarities. But his subtle resolution underestimates the lasting power of Learned Hand's dictum, and other courts have not always been so careful. In particular, consideration of dissimilarities is often decisive in works with thin copyrights.²⁰⁸ In such cases, the original and accused works are likely to be perceived as similar in the colloquial sense. However, that similarity may derive principally from unprotectable elements, such as material dictated by functional requirements or taken from the public domain. Being able to rely on differences to rebut such perceived similarity thus becomes quite important in these cases.

206. *Id.* at 241 (citations omitted).

207. *Id.* at 243–45.

208. *See supra* Part I.C.3 (explaining notion of thin copyright).

Tufenkian Import/Export Ventures, Inc. v. Einstein Moomjy, Inc. provides an example.²⁰⁹ *Tufenkian* was a dispute between two rug designers. The plaintiff, James Tufenkian, developed his “Floral Heriz” design by combining two rugs in the public domain, the “Battilossi” (a Persian antique) and the “Blau” carpet (an Indian Agra).²¹⁰ The principal field of the carpet was taken from the Battilossi; Tufenkian elongated the design, selectively culled some motifs from the dense Battilossi pattern, and placed it slightly off center to create an asymmetrical effect.²¹¹ Tufenkian took the main border from the Blau, and added original minor borders.²¹² The defendant’s rug design combined the same two public domain sources—Battilossi for the field and Blau for the border. However, the accused work used the raw components differently, adding a second “beetle” element to the field, employing a symmetrical design, modifying the main Blau border in a different way, and making different selections in removing motifs from the Battilossi—keeping one the plaintiff did not, and removing one that the plaintiff kept.²¹³

The district court found that Tufenkian had a valid copyright in his rug design, though it offered only thin protection “due to the very substantial incorporation of public domain elements.”²¹⁴ Accordingly, the judge applied the more discerning observer test.²¹⁵ He found that the concept of combining the Battilossi and Blau public domain rugs was an “idea,” and discounted similarity based on the overall design.²¹⁶ On this view, the rug’s protectable elements were limited to the “removal of certain elements to create open space, the asymmetrical pattern, the elongation of the design adapted from the body of the Battilossi rug, the adaptation of the Blau rug border, [and] the creation of the castle and stick figure animal borders.”²¹⁷ These elements, however, were not copied—the defendant’s design was symmetrical, it selected different elements to remove, and it modified the Blau border in a different way.²¹⁸ Recognizing that “any

209. 338 F.3d 127 (2d Cir. 2003).

210. *Id.* at 129.

211. *Id.*

212. *Id.* at 129–30.

213. *Id.* at 130. Pictures of the original and accused rugs, as well as the Battilossi and Blau, can be found in the appendices to the district court opinion. See *Tufenkian Import/Export Ventures, Inc. v. Einstein Moomjy, Inc. (Tufenkian District Court)*, 237 F. Supp. 2d 376, 390–393 (S.D.N.Y. 2002), vacated, 338 F.3d 127 (2d Cir. 2003).

214. *Tufenkian District Court*, 237 F. Supp. 2d at 384.

215. *Id.* at 386.

216. *Id.* at 388.

217. *Id.* at 387.

218. *Id.* at 388.

written description cannot adequately capture the totality of the comparison,” the judge noted that, at his request, the parties brought the full-sized rugs into court so that he could personally view them.²¹⁹ Based on this analysis, he granted summary judgment to the defendants.²²⁰

The Second Circuit reversed, holding that the rugs were substantially similar as a matter of law.²²¹ The appeals court agreed that the more discerning observer test applied, but it gave far more weight to similarity in “total concept and feel.”²²² Some of the difference between the courts lies in the fact that the Second Circuit judged the defendant’s selective removal of motifs to be more similar to the plaintiff’s than the district court did.²²³ However, the Second Circuit also held that the defendant’s addition of a new element—the “beetle” theme—was of no import to the analysis. Citing *Sheldon*, the court reasoned that “[t]o hold otherwise would be rather like holding that one who closely copies a long poem can do so without prima facie infringement if she replaces one (admittedly significant) stanza with verse of her own invention.”²²⁴ Thus, even in the Second Circuit—and in a case involving a graphical work—not all courts recognize Judge Newman’s admonishment that differences can act to undercut similarities.

Kurt S. Adler, Inc. v. World Bazaars, Inc. is another thin copyright case where consideration of differences effectively determines the outcome.²²⁵ *Adler* presented the question of whether two Christmas tree ornaments, each depicting a Santa Claus who was equipped to blow soap bubbles, were substantially similar.²²⁶ As the district court recognized, *Adler* is a difficult case because so little of each work is protectable expression. First, there was no dispute that “stereotypical elements” of Santa Claus—“a jolly, rotund, elder gentleman, wearing a red suit and floppy cap with white trim, and a black belt and boots”—are unprotectable.²²⁷ Moreover, the elements of the work that enabled the Santas to blow bubbles—“a

219. *Id.*

220. *Id.* at 389.

221. *Tufenkian Import/Export Ventures, Inc. v. Einstein Moomjy, Inc.*, 338 F.3d 127, 129, 137 (2d Cir. 2003). This unusual procedural posture is further evidence of the confusion in substantial similarity—both the district and appellate court felt certain enough in their assessment to grant judgment as a matter of law, but to opposite parties.

222. *See id.* at 133–34; *see also supra* Part I.B.3 (describing the total concept and feel test).

223. *Tufenkian*, 338 F.3d at 135–36. In a footnote, the Second Circuit suggested, but did not decide, “that the district court probably erred in categorically dismissing as an unprotected ‘idea’ the plaintiff’s combination of the Battilossi half-field and Blau-ish border.” *Id.* at 135 n.11.

224. *Id.* at 137.

225. 897 F. Supp. 92 (S.D.N.Y. 1995).

226. *Id.* at 93–94.

227. *Id.* at 95.

pivoting arm holding a bubble wand, which arm is capable of dipping the wand into a reservoir of bubble fluid and bringing the wand to a round hole in Santa's mouth"—are functional elements that must also be excluded from the similarity analysis.²²⁸

Nonetheless, the district court found infringement was likely. The judge focused on similarity in "overall appearance," noting that each Santa had "a pear shaped head, a red underlip emphasized, an upcurving mustache, a skin tone bubble nose, rounded boots, [and] a similarly shaded green basin."²²⁹ However, the two Santas were different in many other ways: one's eyes had "crowfeet"; the beards were different; one had black gloves, the other green mittens; and the postures and bases were different.²³⁰ Although the judge in *Adler* did not adhere to the *Sheldon* rule and completely disregarded these dissimilarities, he gave them very little weight. *Adler* held that the differences were "all but irrelevant to the overall appearance" because an ordinary observer would overlook them.²³¹ Of course, most of the similarity in overall appearance was doubtless due to the fact that both toys depicted a stereotypical Santa Claus.

In sum, courts give varying degrees of weight to dissimilarities in the substantial similarity analysis. Strict application of the *Sheldon* rule suggests that they are irrelevant, and courts such as the appeals court in *Tufenkian* take this at face value, ignoring additions and changes in the accused work. By contrast, dissimilarities were at the heart of the case in *Warner Bros.*, undercutting quite numerous similarities. Finally, courts such as the district court in *Adler* take an intermediate approach, formally considering dissimilarities but affording them little weight, even in a case involving a thin copyright.

III. A PROPOSED CONTEXTUAL APPROACH

This Part uses the case law examples discussed above to offer some proposed resolutions to the framing problems in similarity analysis. In general, I advocate for a more contextual, flexible, and subtle approach to these vexing issues than courts have usually taken. Such an approach strives to put the defendant's copying in context, viewed against both the original work and the accused work. It considers not only the perspective of the original creator, but of the alleged infringer as well. It rejects

228. *Id.*

229. *Id.*

230. *Id.*

231. *Id.* at 95–96.

selecting only the original or accused work as the sole baseline in similarity analysis, or disregarding important differences or the addition of new material. It endeavors not to dismiss relevant information about what was copied, how it was used, and why.

This proposed approach eschews formalist rules, like the *Sheldon* rule, that result in courts focusing only on the original work as the baseline, and stripping the accused work of its context. For example, why ignore subsequent events—such as developments that made copyrightable expression a functional standard—when those developments were the very reason the copying was made? Why disregard how the copied expression was used in the context of the defendant’s work when assessing its substantiality? The path detailed below seeks to assume the perspectives of both the original creator and alleged infringer, as relevant, in attempting to draw the elusive line past which copying becomes significant enough to be actionable.

It should be made clear at the outset that nothing in these proposed resolutions encroaches upon or diminishes that *other* critical limitation on the scope of copyright, fair use. Fair use is independent of substantial similarity, arising only after the plaintiff has proven infringement.²³² While there may be some overlap between the two doctrines—both consider, for example, how much expression was copied²³³—they remain distinct and independent. Fair use considers many factors, such as educational or parodic purpose, that even the most broadly conceived similarity analysis would not.²³⁴ The suggested reforms to substantial similarity outlined below are thus intended only to complement, not to supersede, fair use. Substantial similarity may receive less attention, but it is at least as complex as fair use, and of more general application.²³⁵ A significant motivation of this Part is to ensure that substantial similarity, like fair use, remains a robust and vital limitation on copyright.

232. See Balganes, *supra* note 40, at 206 (citations omitted) (“[Substantial similarity] bears no connection whatsoever to the fair-use doctrine, a doctrine which operates as a defense to infringement rather than as a component of the plaintiff’s prima facie case.”).

233. Compare 17 U.S.C. § 107(3) (2014) (“[Fair use considers] the amount and substantiality of the portion used in relation to the copyrighted work as a whole . . .”), with *Ringgold v. Black Entm’t Television, Inc.*, 126 F.3d 70, 75 (2d Cir. 1997) (“[Substantial similarity] requires that the copying is quantitatively and qualitatively sufficient to support the legal conclusion that infringement (actionable copying) has occurred.”).

234. See, e.g., *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579–84 (1994) (giving heavy weight to a work’s parodic elements under the first fair use factor).

235. See Balganes et al., *supra* note 2, at 268 (“[T]he complexities of the fair use doctrine pale in comparison to what is central to almost all cases of copyright infringement: the question of ‘substantial similarity.’”).

A. *Applying the Idea-Expression Distinction Dynamically*

This Part argues that the idea-expression distinction should be evaluated dynamically. In particular, subsequent events that turn once-protectable elements into unprotectable ideas or functional standards must be considered in determining whether copyright infringement has occurred. This Part conceives of similarity's timing problem as two related but distinct splits in judicial authority: (1) whether the doctrines of merger and *scènes à faire* are assessed as part of copyrightability or infringement; and (2) whether the idea-expression distinction should apply at the time of creation or the time of infringement. It argues that the second question is the critical issue, and that copyright's constitutional purpose to promote the creation and dissemination of knowledge requires that the idea-expression distinction apply in a dynamic fashion. In effect, this means that material must be protectable by copyright *both* at the time of creation (in order to be original and copyrightable in the first instance) and at the time of infringement (in order to avoid idea-expression merger).

Recall that similarity's timing problem concerns how to handle elements that were once copyrightable expression, but subsequently become or merge with an idea that is not protectable by copyright.²³⁶ Courts such as the district court in *Lexmark* focus on the perspective of the original creator, and conclude that the work is protectable by copyright if it could have been expressed in multiple ways at the time of creation.²³⁷ The Federal Circuit in *Oracle* followed similar logic, focusing on the options available at the time of creation, and ignoring the interoperability concerns that motivated Google's copying.²³⁸ In contrast, the appeals court in *Lexmark* and the district court in *Oracle* looked at the choices available to the alleged infringer at the time of infringement, ruling for the defendants because what was copied had become a functional standard or method of operation.²³⁹

These varying approaches derive from two distinct splits in authority, and much confusion has resulted from courts' failure to be precise about the difference. The first issue—the primary focus of the scholarly literature—is whether merger and *scènes à faire* should be part of the

236. See *supra* Part II.A (explaining and providing examples of the timing problem). Although the discussion here principally uses “idea” as a shorthand for any non-expressive elements, the same analysis of course applies to other unprotected categories such as a “procedure, process, system, method of operation, concept, principle, or discovery.” 17 U.S.C. § 102(b) (2014).

237. See *supra* notes 140–42 and accompanying text.

238. See *supra* notes 125–29 and accompanying text.

239. See *supra* notes 123–27, 146–49 and accompanying text.

copyrightability inquiry or part of the infringement inquiry.²⁴⁰ At minimum, the initial copyrightability inquiry includes the requirement that the plaintiff's work be *original*; that is, independently created and possessing "at least some minimal degree of creativity."²⁴¹ Originality thus requires some modicum of artistic choice in creation, and those choices must not be commonplace, mundane, or dictated by efficiency or functional concerns—or else the work is not copyrightable at all.²⁴² Some courts go further to apply the merger doctrine, too, as part of this copyrightability analysis.²⁴³ Other courts view merger as part of the substantial similarity/infringement analysis.²⁴⁴

The second split is when to apply the idea-expression distinction. This issue becomes critical when an expressive element becomes, with time, an idea, system, functional standard, or method of operation. In other words, when determining what is an "idea," should we look to the choices and constraints facing the original author at the time of creation, or the alleged infringer at the time of infringement? The competing opinions in *Oracle* illustrate this divide. The district court focused on the choices available to Google at the time of copying, finding merger because Google copied only what was necessary to achieve a degree of interoperability.²⁴⁵ The Federal Circuit focused instead on the time of creation, concluding that there was no merger and dismissing Google's claim that the API names had become "industry standard[s]."²⁴⁶

While one might presume that these two issues are rigidly linked—that if merger is part of copyrightability, one should define the scope of "idea" at the time of creation—they are, in practice, independent. Judge Sutton's opinion in *Lexmark*, for example, takes what we might call a "mixed" view. Though the opinion is not perfectly clear on this point, Judge Sutton's holding focuses primarily on the Toner Loading Program's lack

240. See, e.g., 2 PATRY, *supra* note 75, § 4:46 ("[M]erger, if applied at all, should be applied at the infringement, not at the originality stage of analysis."); Ocasio, *supra* note 150, at 310–23 (arguing that merger should apply as part of copyrightability); Teter, *supra* note 111, at 1075–77 (arguing that merger should be applied "as a question of substantial similarity rather than as a question of copyrightability").

241. *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 345 (1991).

242. See *id.*; see also *Matthew Bender & Co. v. W. Publ'g Co.*, 158 F.3d 674, 681–83 (2d Cir. 1998).

243. See *supra* Part I.C.2 (explaining merger, *scènes à faire*, and functionality doctrines).

244. See, e.g., *Lexmark Int'l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522, 557–59 (6th Cir. 2004) (Feikens, J., concurring in part and dissenting in part) (explaining and citing cases on both sides of this split); Ocasio, *supra* note 150, at 310–23 (reviewing circuit split).

245. *Oracle Am., Inc. v. Google Inc. (Oracle District Court)*, 872 F. Supp. 2d 974, 1000–01 (N.D. Cal. 2012), *rev'd*, 750 F.3d 1339 (Fed. Cir. 2014).

246. *Oracle Am., Inc. v. Google Inc.*, 750 F.3d 1339, 1361–62, 1372 (Fed. Cir. 2014).

of originality, both because the TLP's structure was dictated by functionality *and* because Lexmark used it as a "lock-out code" for their printers.²⁴⁷ In effect, the opinion treats merger as part of the copyrightability/originality analysis, but it still applies idea-expression at the time of infringement because it considers the TLP's subsequent use as a lock-out code. On the other side, the Federal Circuit in *Oracle* purports to apply merger as part of the infringement analysis, but nonetheless applies the idea-expression distinction at the time of creation, focusing only on the choices available to Oracle.²⁴⁸ The chart below summarizes the approaches taken in the cases discussed in Part II.A.

TABLE 1: DIFFERING JUDICIAL APPROACHES TO THE TIMING PROBLEM

	<i>When the Idea-Expression Distinction Is Applied</i>		
		<u>Time of Creation</u>	<u>Time of Infringement</u>
<i>Stage at Which Merger and Scènes à Faire Are Assessed</i>	<u>Copyrightability</u>	<i>Lexmark</i> (E.D. Ky.) <i>Result: Plaintiff wins</i>	<i>Lexmark</i> (Sutton); <i>Oracle</i> (N.D. Cal) <i>Result: Plaintiff loses, invalid copyright</i>
	<u>Infringement</u>	<i>Oracle</i> (Fed. Cir.) <i>Result: Plaintiff wins</i>	<i>Lexmark</i> (Feikens) <i>Result: Plaintiff loses, no infringement</i>

Though much of the literature and case law focuses on the first issue, the second split is the critical one. The copyrightability versus infringement question, as a practical matter, only affects remedy—whether the plaintiff loses his copyright entirely (no copyrightability), or merely loses the particular dispute (no infringement). It is the timing issue that determines whether the plaintiff wins or loses. When courts look only to the time of creation, they tend to conclude that the plaintiff's work involved creative choice that was not dictated entirely by functional concerns, and thus find liability. If courts instead look to the choices available to the infringer, they usually conclude that the allegedly

247. See *Lexmark*, 387 F.3d at 539–41.

248. See *Oracle*, 750 F.3d at 1358.

infringing expression became (or “merged” with) a functional idea, and find no liability. So long as a court applies the idea-expression dynamically, the scope of the plaintiff’s loss (no copyright versus no infringement) is of less import. In the usual case, a finding of no infringement will effectively make the copyright unenforceable, even if the ruling was formally limited to infringement. The critical feature of these disputes is whether “idea” is defined flexibly, to recognize that subsequent developments and context can transform seeming expression into an unprotectable idea, procedure, or method of operation.

Lexmark provides the clearest example.²⁴⁹ Stripped of context, the TLP may contain some protectable expression (if not a lot)—it is just a simple program that checks a toner level that, we will assume *arguendo*, could have been written in multiple ways. At the time of creation, it may have contained copyrightable expression. But it is abundantly clear that, as Lexmark used it and as SCC copied it, the TLP cannot be protected. To do so would be to ignore the context that transformed the TLP into a lock-out code—a method of operation that it was necessary to copy in order to make use of the printer. As Judge Feikens’ example made clear, it would not matter whether the TLP was undoubtedly expressive material (even a poem) out of context. What matters is that the defendant was only copying it for the functional purpose of operating the printer. Thus, as used, the TLP cannot be protectable even if it were copyrightable at the time of creation.

It is difficult to defend decisions like the Federal Circuit’s in *Oracle* that take a contrary approach. There is simply no good legal or policy reason to employ an artificially constrained definition of “idea.” Looking only at the time of creation ignores subsequent events and context that are often the heart of why the alleged infringer copied. The highly formalistic approach of the Federal Circuit in *Oracle* denies the practical reality of the constraints faced by Google due to context and circumstances. It is also inconsistent with the copyright statute, which states that “[i]n no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery”²⁵⁰ It is hard to imagine language that is more absolute and unqualified. Copyright simply cannot be applied to ideas, methods of operation, and the like—regardless of *when* they become such.

249. See *supra* notes 131–49 and accompanying text.

250. 17 U.S.C. §102(b) (2014) (emphasis added).

More fundamentally, the Federal Circuit's approach in *Oracle* threatens the public's right to the free flow of ideas, the balance at the core of copyright.²⁵¹ It goes against the core constitutional purpose of "promot[ing] . . . Progress"²⁵² for copyright to lock up something as general as an idea, regardless of *when* it attained that status. Copyright's extraordinary term of the author's life plus seventy years makes it particularly absurd to ignore changed circumstances and dynamic effects.²⁵³ This is an especially acute problem in the context of software, where material can quickly become a widely adopted industry standard.²⁵⁴ The policy case for a dynamic idea-expression distinction is really quite overwhelming.²⁵⁵ In the most generous view, it may be that courts' view of the first issue, whether merger is part of copyrightability or part of infringement analysis—which is debatable—is clouding their view of the second, more fundamental timing concern.

That leaves the issue of remedy: if idea and expression merge subsequent to creation, should we deny copyrightability altogether, or simply find no infringement? Contrary to the views of some commentators, a holding that the copyright is entirely invalid may go too far.²⁵⁶ Judge Feikens' example of a poem used as a lock-out code illustrates that a work may be an unprotectable idea in one context, but protectable expression in another. Thus, viewing merger as part of infringement, but evaluating the idea-expression distinction at the time of infringement, offers a reasonable middle ground. The alternative approach—taken by Judge Sutton in *Lexmark*, for example—is oddly

251. See *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 349–50 (1991) (citing *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556–57 (1985)) ("[C]opyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work.").

252. U.S. CONST. art. I, § 8, cl. 8.

253. 17 U.S.C. § 302(a) (2014) (setting general copyright term as the life of the author plus seventy years).

254. See generally sources cited *supra* note 151.

255. See, e.g., Menell, *supra* note 111, at 1066–68, 1099–1101 (arguing that network effects urge application of merger doctrine to prevent protection of a "de facto industry standard"); Samuelson, *supra* note 111, at 196–215, 221–24 (arguing that coding standards should be uncopyrightable under 17 U.S.C. § 102(b) due to competition and public policy concerns); Fred Anthony Rowley, Jr., Note, *Dynamic Copyright Law: Its Problems and a Possible Solution*, 11 HARV. J.L. & TECH. 481, 495–96 (1998) (arguing that copying of standard user interface elements should be permitted to avoid locking out subsequent innovators); Teter, *supra* note 111, at 1066–72 (arguing that copying to permit compatibility must be permitted as a matter of policy).

256. See, e.g., Ocasio, *supra* note 150, at 310–23 (arguing that merger should be applied as part of the copyrightability analysis); Preonas, *supra* note 111, at 97–109 (arguing that, in the context of factual compilations, "mergercide" should invalidate copyright when once-original expression becomes an idea).

achronological, as it applies merger as formally part of copyrightability but still looks ahead to time of infringement. This gets the important issues right, but it is more conceptually coherent to apply merger at the time of infringement and as part of the infringement analysis. This also avoids the perceived harshness of a complete forfeiture of copyright, while still permitting copying for purposes of functionality or compatibility. In effect, such a view adopts the perspective of both the original creator and alleged infringer. It looks to the choices available to the original author at the time of creation as part of the originality analysis, but also considers the options available to the alleged infringer at the time of infringement, when considering substantial similarity.

B. A Qualitative, Contextual Approach to Substantiality and Dissimilarities

This Part argues against using either only the plaintiff's work or the defendant's work as the baseline for assessing substantiality. The quantitative, percentage-style baselines invoked in cases such as *Harper & Row* and *Newton v. Diamond* needlessly ignore important context—either the nature of the use, or the nature of the original work.²⁵⁷ Indeed, outside of the context of linear, literary works, the concept of quantitative measure becomes incoherent.

For similar reasons, courts should not ignore important dissimilarities between the original and accused works by focusing solely on the plaintiff's work. The tendency of some courts to disregard important dissimilarities results from a misunderstanding of the *Sheldon* rule and its misapplication far beyond its original purpose. Instead, courts should take a primarily qualitative approach to substantiality, considering relative importance contextually, using both works as baselines and not dismissing dissimilarities as irrelevant to the infringement analysis.

The current approach to the baseline and dissimilarity problems—to the extent there is any consensus at all—is embodied in the *Sheldon* rule. Recall that *Sheldon* instructs courts to ignore the changes that an alleged infringer has made because “no plagiarist can excuse the wrong by showing how much of his work he did not pirate.”²⁵⁸ *Sheldon* leads courts to take the plaintiff's work as *the* baseline for comparison, which has the effect of making the original work the more prominent, and biases the

257. See *supra* Part II.B (reviewing case law examples of the baseline problem).

258. See *supra* notes 176–79 and accompanying text.

comparison in favor of the plaintiff.²⁵⁹ (This concern is exacerbated when the plaintiff's copyrighted work is well known, a not uncommon scenario in copyright disputes.²⁶⁰) Many courts invoke *Sheldon* to justify using the plaintiff's work as the baseline when measuring the quantitative significance of the defendant's copying.²⁶¹ Some courts have extended the *Sheldon* rule to disregard dissimilarities between the two works and changes that the defendant has made.²⁶²

The *Sheldon* rule is triply wrong. First, the "tacking on" concern that motivated *Sheldon* is overstated, and courts have not been careful to limit the *Sheldon* rule to its original context of linear, literary works. Second, a strict application of *Sheldon* causes courts to ignore the context of the defendant's use, context that may be critical to understanding whether the alleged infringement should be actionable. *Sheldon* is thus unfair to defendants when it ignores the new material that they added and the context in which the alleged infringement was made. Finally, *Sheldon* is incorrect to focus myopically on the plaintiff's work as the baseline. This can result in unfairness to the plaintiff, such as in the *Antonick* case, when larger works are effectively penalized for involving additional expression. Courts' tendency to shift baselines in an ad hoc fashion may well be an indication of the tensions created by a strict application of *Sheldon*.

The concern that motivated *Sheldon* is the "tacking on" problem—the idea that an infringer could escape liability simply by adding a mass of new, extraneous material to his infringing work.²⁶³ This concern is overstated. It is not apparent in the facts of *Sheldon* itself, and it is far from clear that the tacking-on problem is a realistic scenario. In any event, even presuming the concern is real, there are better and more direct ways to deal with the problem of "manipulating" the de minimis inquiry.²⁶⁴ Blatant attempts to avoid infringement by tacking on additional material can be easily ferreted out directly by courts, and dismissed via equitable doctrines.²⁶⁵

259. See *supra* notes 13, 108–09 and accompanying text (summarizing results from cognitive science that greater similarity is perceived when the more prominent of two concepts is treated as the baseline).

260. See, e.g., *Warner Bros. Entm't Inc. v. RDR Books*, 575 F. Supp. 2d 513, 518 (S.D.N.Y. 2008) (finding infringement of *Harry Potter* series of novels); *Castle Rock Entm't, Inc. v. Carol Publ'g Grp., Inc.*, 150 F.3d 132, 135 (2d Cir. 1998) (finding infringement of *Seinfeld* television series).

261. See *supra* notes 164–75 and accompanying text (discussing *Newton v. Diamond*).

262. See *supra* notes 209–24 and accompanying text (discussing *Tufenkian*).

263. See *supra* notes 179–80 and accompanying text.

264. Cf. 3 PATRY, *supra* note 5, § 9:65 (“[*Sheldon*] has rightly been cited as authority against attempts to manipulate the de minimis [infringement] inquiry.”).

265. See, e.g., *Petrella v. Metro-Goldwyn-Mayer, Inc.*, 134 S. Ct. 1962, 1977 (2014) (indicating

In the more usual case, what was added by the defendant is not “tacked on,” but rather important, and should not be ignored. The addition of new expression by the defendant is typically not extraneous, unrelated material—as envisioned by *Sheldon*—but instead precisely the type of expression one might *want* to encourage through copyright. Take *Newton v. Diamond* as an example.²⁶⁶ Though one could view the rapping and music that the Beastie Boys added as “tacked on,” that is a ridiculous claim—the material added is the heart of how the Beastie Boys used the sample to make a new work. Notably, the *Sheldon* rule is precisely the opposite of what courts do in the context of fair use, where much of the emphasis is rightly on whether the defendant transformed the original work and added “something new.”²⁶⁷ In the context of similarity, the defendant’s new material is relevant both to the perceived similarity of a lay observer, as well as to any economic harm caused by the infringement.²⁶⁸

It is therefore odd to dismiss the defendant’s new material or dissimilarities between the two works out of hand. Why would we *not* consider the defendant’s actual use in the context of his new work? In *Newton*, the Beastie Boys should certainly prevail, but not because their looping is “irrelevant.”²⁶⁹ Of course it’s relevant. If “Pass the Mic” were nothing but Newton’s flute sample looped for three minutes—no rapping, no added samples and instrumental music—would it not be a different, and tougher, case? Moreover, although ignoring the defendant’s use in *Newton* worked to the advantage of the defendant, usually it is used in the opposite manner, as it was in *Sheldon* itself. This is true in the *Tufenkian* case, for example, where *Sheldon* was invoked to prevent the defendant from relying on dissimilarities to negate the infringement claim.²⁷⁰ In other words, although *Sheldon* saves the Beastie Boys from having to explain looping, it leaves them unable to rely on the fact that they added new samples, rapping, and live instruments over the Newton sample. Similarly,

that equitable doctrine of estoppel is available in copyright actions); *Saxon v. Blann*, 968 F.2d 676, 680 (8th Cir. 1992) (confirming that the doctrine of unclean hands is applicable in copyright actions).

266. See *supra* notes 164–75 and accompanying text (discussing *Newton*).

267. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994).

268. See Jeanne C. Fromer & Mark A. Lemley, *The Audience in Intellectual Property Infringement*, 112 MICH. L. REV. 1251, 1255 (2014) (arguing that infringement in intellectual property regimes should only be found when works are similar from a technical perspective, and cause market harm from the view of consumers); Lape, *supra* note 10, at 202 (arguing that substantial similarity requires a finding of “economic harm” or other “injury that may have an impact on authors’ incentive to create”).

269. *Newton v. Diamond*, 388 F.3d 1189, 1195 (9th Cir. 2003).

270. See *supra* notes 209–24 and accompanying text (discussing *Tufenkian*).

it leaves the defendants in cases like *Tufenkian* or *Adler* unable to rely upon the important dissimilarities between their works and the original works. Considering only what the defendant copied, without considering how the defendant used what was copied in context, or the changes she made, is a form of willful blindness that tends to bias the inquiry against defendants.

The *Sheldon* rule can sometimes be unfair to plaintiffs as well. In particular, using the plaintiff's work as the measure for quantitative substantiality effectively penalizes the plaintiff for creating more expression. Consider the facts of the *Antonick* case, for example.²⁷¹ There, no substantial similarity was found because the original work was complex—a football simulation video game—even though EA had copied Antonick's code for plays and formations. When the original work happens to be large—be it Ford's ponderous memoir in *Harper & Row*, or the millions of lines of code in *Antonick* or *Oracle*—percentage-style measures using the plaintiff's work as the baseline dilute the significance of what was taken. In *Harper & Row*, three hundred words is a trivial percentage of Ford's memoir, but it might still be substantial in context. As Table 2 shows, percentage-based measures can vary widely depending on the baseline.

TABLE 2: QUANTITATIVE MEASURES OF HOW MUCH WAS COPIED USING VARIOUS BASELINES

Baseline	<i>Harper & Row</i>	<i>Newton v. Diamond</i>	<i>Oracle v. Google</i>	<i>Antonick v. Electronic Arts</i>
<u>Plaintiff's work</u>	0.1% ("insubstantial")	2% of running time; greater % of score	3% of copied APIs; trivial % of Java	small
<u>Defendant's work</u>	"at least 13%"	looped background element; significant (1/5) part of song	trivial % of Android	small
<u>Absolute taking</u>	300 words, plus paraphrasing	six seconds; three note phrase	7,000 lines of code	code for plays and formations

271. See *supra* notes 182–92 and accompanying text (discussing the *Antonick* case).

Using the plaintiff's work as *the* baseline thus seems perverse if copyright is designed to incent new expression. Taken literally, it effectively penalizes original creators for creating more expression. The real issue in these cases is whether what was taken was substantial *in itself*, and qualitatively significant in the context of both the original and accused works. The court in *Antonick* should consider whether the plays and formation code is significant in overall qualitative importance to the games, how many lines of code total were copied, and how much of that was commonplace or dictated by functionality. To dismiss the plays and formation code as insignificant only because it is small piece of a larger game misses the point. In this respect, Justice O'Connor's focus on the qualitative import of the expression in *Harper & Row* was correct. However, the sleight of hand involved in invoking the defendant's work as the baseline in *Harper & Row*—"at least 13% of the infringing article"—is misleading.²⁷² These percentage-style measures are red herrings. Would *Harper & Row* really be a different case if Ford's memoir were only 10,000 words, or if *The Nation's* article were twice as long? Does it really matter in *Oracle* how many millions of lines of code are in Java or Android? Probably not, and courts should not imply otherwise.

Furthermore, the notion of a percentage measure is unsound for non-linear works. In literary works, we can at least count the number of words or lines of code taken—though, of course, paraphrasing may present some difficulties. For musical works, one instinct is to look to the percentage of the running time, as the *Newton* court did. But it is not clear why that measure is superior to using the number of notes, the percentage of the score, or the prominence of a sound in the mix. Running time also becomes deceptive when there is more than one sample or instrument playing simultaneously, as is usual. In theory, one could try to apportion this—say, 25% of two minutes for the first sample, etc.—but this quickly becomes arbitrary. More importantly, such quantitative measures only avoid the real issues of qualitative significance. In *Newton*, the courts should be asking whether the idea of singing into a flute is even protectable expression at all, not whether "six seconds" or "three notes" is a significant part of Newton's composition. If Newton had squeezed a Charlie Parker-style flurry of notes into the six-second sample, or if "Choir" was only twenty seconds long, the measures the majority relied upon in *Newton* would greatly change. However, the fundamentals of the case would not be truly different.

272. *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 565–66 (1985).

The difficulties of quantitative measurement are even more serious for pictorial, graphic, and sculptural works.²⁷³ For visual works intended to be perceived as an entirety, one can scarcely begin to imagine how to define a percentage appropriation. Take the two bubble-blowing Santa ornaments in *Adler* as an example.²⁷⁴ How does one calculate what “percentage” of the ornaments are the same, and how much of that was due to protectable elements? We are at a loss to say. In theory, perhaps, one could analyze a two-dimensional work pixel-by-pixel and count the pixels that are identical. But this method is coherent, and then just barely, only in the special case of pastiche.²⁷⁵ A pixel-by-pixel analysis would likely yield zero or a trivial number of similarities in a case where an artist’s visual style was copied—*Steinberg v. Columbia Pictures Industries* is a well-known example—even though it is clear to a lay observer that something was taken (whether we think that something is protectable or not).²⁷⁶ This pixel-based measure also runs aground without a way to exclude all unprotected material. The rugs in *Tufenkian*, for example, might be highly similar on pixel percentage basis, but that is likely only because they are both derived from the same public domain sources.

In sum, courts should focus mainly on the contextual, qualitative importance of what was copied in assessing its significance in substantial similarity. If the *Sheldon* rule must be invoked at all, it should be limited to its original purpose—cases where a plaintiff or defendant attempts to manipulate the infringement inquiry by adding extraneous material. To the extent courts must consider the copying quantitatively at all, the best course may be to measure the expression copied as an absolute matter. Such a measure looks to the total number of words copied, the number of lines of code, or the length of a piece of music, but always in light of the qualitative significance of that expression in context, including its level of originality and its economic significance.²⁷⁷ The absolute measure at least gets courts thinking about the right questions—for example, whether what was copied was too short, too commonplace, or too functional to be

273. See 17 U.S.C. § 102(a)(5) (2014).

274. See *supra* notes 225–31 and accompanying text (discussing *Adler*).

275. See *Cariou v. Prince*, 714 F.3d 694, 705–12 (2d Cir. 2013) (assessing whether series of appropriation art pastiche of photographs from earlier works was a fair use).

276. 663 F. Supp. 706 (S.D.N.Y. 1987). That case involved a poster for the movie “Moscow on the Hudson” that evoked the style of Saul Steinberg, an artist for the *New Yorker*, and in particular his famous work “View of the World from 9th Avenue.” *Id.* at 708–11.

277. Cf. Busek, *supra* note 10, at 1795–1803 (arguing for an approach to substantial similarity based on the degree of possible expressive variation); Lape, *supra* note 10, at 194–206 (arguing for an approach to substantial similarity based on substantial economic or other harm to the plaintiff).

protectable expression—and not its *relative* shortness vis-à-vis the plaintiff's or defendant's work.²⁷⁸ Similarly, courts should not disregard material added by the defendant, or important dissimilarities that undermine perceived similarity. Ignoring dissimilarities is another unintended result of *Sheldon's* dictum. It is a mistake to adhere to this rule literally, as the Second Circuit did in *Tufenkian*. The balance struck by Judge Newman in *Warner Bros.* is a sounder approach.

CONCLUSION

This Article has sought to bring attention to the framing problems that plague courts' analyses of similarity in copyright infringement. Courts take a distressingly inconsistent approach to the issues of baselines, timing, perspective, and dissimilarities when structuring the comparison between the original and the accused works. The result is widespread confusion, unpredictability, and contradictory results. This Article argues that courts should take a flexible, contextual approach to similarity's framing problems. Such an approach is fairer to both parties, more predictable, and truer to copyright's purpose of promoting the creation and dissemination of knowledge. It avoids the manipulation and biases created by the more rigid approaches that predominate in the case law. Most importantly, it serves to guarantee substantial similarity's role as a robust and vital limitation on the scope of copyright.

278. This is one of the few things that the Federal Circuit got right in *Oracle*. The fact that “7,000 lines of code” were copied is probably more relevant than the fact that those lines are only a trivial aspect of Java or Android. *See Oracle Am., Inc. v. Google Inc.*, 750 F.3d 1339, 1356 (Fed. Cir. 2014). However, the qualitative significance of the code in that case makes it obvious that no infringement should be found. The key issue in that case was not the total amount taken, but its standard, commonplace, and functional nature. Arguably, Google might be permitted to copy the whole of the APIs.