2019

Redundancy and Anti-Redundancy in Copyright

Oren Bracha
John M. Golden

Follow this and additional works at: https://opencommons.uconn.edu/law_review

Recommended Citation
https://opencommons.uconn.edu/law_review/414
Article

Redundancy and Anti-Redundancy in Copyright

OREN BRACHA & JOHN M. GOLDEN

United States copyright law has a reputation as a tangled mess of overlapping legal doctrines. Although commonly malign ed, this redundancy can play a positive role. Redundancy that is well-designed and implemented can achieve a better balance between copyright law’s benefits and costs, safeguard interests in competition and technological innovation, improve protection of freedom of speech, and enable flexibility and doctrinal evolution. On the other hand, doctrinal redundancy can have unfortunate results when it is excessive or otherwise unmoored from underlying purpose. Design principles of redundancy and anti-redundancy are thus virtually tailor-made to analyze copyright law’s structure in search of identifying potential justifications, flaws, and opportunities for reform. Building on prior work on legal redundancy and on copyright doctrine, this Article examines the promise and risks of doctrinal redundancy as a design principle for copyright. Specifically, the Article analyzes redundancy in four areas of copyright doctrine: (1) fair use and non-functionality; (2) fair use and improper appropriation; (3) copyrightable subject matter and originality; and (4) the reproduction and derivative-work rights. The analysis indicates that distinct but overlapping doctrines of fair use, originality, and copyrightable subject matter can better police the boundaries of an expansive copyright regime than could a single doctrine alone. Such reinforced policing is particularly important to secure interests in free speech and competition against improper copyright encroachment. In contrast, however, there is reason to question the social value of redundancy as embodied in distinctly identified rights in making reproductions and derivative works. More generally, our analysis illustrates how attention to the structural design of legal doctrine can help to improve the content and operation of law.
ARTICLE CONTENTS

INTRODUCTION .................................................................................................................. 249

I. REDUNDANCY AND ANTI-REDUNDANCY OVERVIEW ........................................... 251
   A. FORMS OF REDUNDANCY ....................................................................................... 251
   B. ARGUMENTS FOR REDUNDANCY AND ANTI-REDUNDANCY ............................ 253

II. DOCTORAL REDUNDANCY IN COPYRIGHT ............................................................. 256
   A. FAIR USE AND FUNCTIONALITY ............................................................................ 256
   B. FAIR USE AND IMPROPER APPROPRIATION ....................................................... 273
   C. COPYRIGHTABLE SUBJECT MATTER AND ORIGINALITY ................................. 285
   D. REPRODUCTION AND DERIVATIVE-WORK RIGHTS ......................................... 296

CONCLUSION .................................................................................................................... 304
Redundancy and Anti-Redundancy in Copyright

OREN BRACHA * & JOHN M. GOLDEN **

INTRODUCTION

Recent years have brought new rounds of agitation for substantial copyright reform.1 True reform entails considering a legal field’s central doctrines and working to shape them to best serve their underlying purposes. But reform should commonly do more than consider individual doctrines as separate parts. Like other legal fields, copyright law is a system of legal norms. By “a system” we mean a group of connected norms that mutually interact and affect the results produced by the connected whole. Understanding this systemic dimension goes to the heart of proper evaluation and design of copyright as a legal field. Law is a purposive practice: its norms and procedures are designed to further certain ends.2 Thus, the way that the interaction between different legal norms affect the achievement of underlying purposes is central for the proper design of legal fields.

A rough analogy can be made to problems in systems engineering.3 A good engineer designing a vehicle does not restrict her attention to perfecting each of its elements in isolation. She would also want to know how the

---

* William C. Conner Chair in Law, University of Texas School of Law. For helpful comments, the authors thank David Adelman, Sam Bray, Chris Buccafusco, Janet Freilich, Wendy Gordon, Geoffrey Manne, Peter Menell, Pamela Samuelson, Talha Syed, and participants in a Drawing Board workshop at the University of Texas School of Law, the 2017 Intellectual Property Scholars Conference, the 2018 Fordham IP Conference, and the 2018 Annual Conference of the Society for Institutional and Organizational Economics.

** Loomer Family Professor in Law, University of Texas School of Law.


3 Cf. LON L. FULLER, *The Law in Quest of Itself* 11 (1940) (speaking of “the field of purposive human activity, which includes both steam engines and the law”).
elements interact and affect each other’s function. If the brake system impedes the performance of the otherwise highly effective steering mechanism, this is a pertinent design consideration. Likewise, the engineer will take notice if the combination of two subsystems creates a highly beneficial synergetic effect. If, for some reason, the design of the brake system means that the same desired effectiveness in steering can be achieved with a more streamlined steering mechanism, the engineer might exploit this synergy and use the streamlining to improve another aspect of vehicle performance, such as fuel efficiency.

What can a focus on copyright’s systematic aspect teach us? U.S. copyright law has a reputation of being a tangled mess. Much of this reputation is due to overlapping and sometimes conflicting legal doctrines. Associated redundancies are commonly maligned, but questions of how much overlap or even internal conflict a legal field’s doctrines should feature are more helpfully conceived as questions about appropriate, if not optimal, system design. Redundancy in the structure of legal doctrine can serve positive purposes, such as ensuring effective protection of core interests, providing fail-safes, or enabling flexibility and doctrinal evolution on less-reinforced margins. Of course, not all redundancy is beneficial. Doctrinal redundancy can be wasteful, overly cumbersome for legal decision-makers, or simply confusing. As a result, one can view redundancy and anti-redundancy as competing design principles whose pluses and minuses should be balanced in light of a legal field’s underlying purposes. With its array of overlapping doctrines and need to avoid improper encroachment on interests in free speech and free competition, copyright law is virtually tailor-made for such an analysis. This Article focuses on redundancy and anti-redundancy in examining copyright law as a system of norms.


5 See, e.g., Pamela Samuelson, Reconceptualizing Copyright’s Merger Doctrine, 63 J. COPYRIGHT SOC’Y U.S.A. 417, 453 (2016) (discussing how courts have invoked copyright doctrines that overlap with merger rulings).


7 See id. at 658–66.

8 Id. at 671, 673, 705.

9 See Margaret Jane Radin, A Comment on Information Propertization and its Legal Milieu, 54 CLEV. ST. L. REV. 23, 30 (2006) (“The scheme of copyright law responds to both free speech concerns and competitive concerns by limiting propertization to expression (excluding ideas, facts, functionalities), by having a limited term, by limiting coverage to copying and to distribution of objects that are copies, and by retaining the defense of fair use[.]”).

10 See Bonito Boats v. Thunder Craft Boats, 489 U.S. 141, 150 (1989) (“Taken together, the novelty and nonobviousness requirements express a congressional determination that the purposes behind the Patent Clause are best served by free competition and exploitation of either that which is already available to the public or that which may be readily discerned from publicly available material.”).
In Part I, we provide a brief overview of the nature of doctrinal redundancy and anti-redundancy as well as their advantages and disadvantages. In Part II, we analyze, in light of this theoretical background, various instances of apparent and actual redundancy in copyright law. We argue that such an analysis suggests ways in which the law can be understood, applied, or altered to better serve its social aims. Specifically, we examine four areas of doctrinal overlap: (1) the rule denying copyright protection to functional subject matter and the rule exempting “fair uses” from copyright liability; (2) the improper appropriation requirement and the same fair use doctrine; (3) statutory categories of copyrightable subject matter and the originality requirement; and (4) the intersection of the reproduction and derivative-work rights. We conclude by distilling from these central cases of copyright redundancy what we posit to be general lessons for redundancy and anti-redundancy in legal design.

In multiple cases, we find that what initially might appear to be simply wasteful or superfluous doctrinal repetition is actually—if understood and applied properly—sound legal design that advances important goals, including the protection of core interests in free speech and competition. Indeed, we argue that in some cases, scholarly calls and overt or implied moves by courts to eliminate doctrinal redundancy should be reconsidered. Our outlook, however, is not Panglossian. In one of our cases, we find that a common failure to properly understand the substantive relationship between the two intersecting doctrines has led to confusion and odd results. Moreover, even in the cases where redundancy may have a positive role to play, this is contingent on applying the relevant doctrines with a proper understanding of their interaction and their underlying purpose. In short, we find that copyright law provides both fertile ground for the application of redundancy principles and an enlightening laboratory for understanding them better.

I. REDUNDANCY AND ANTI-REDUNDANCY OVERVIEW

A. Forms of Redundancy

Legal redundancy can appear in procedural, institutional, textual, and doctrinal forms. This Article concerns doctrinal redundancy, which is a form of legal redundancy that occurs when two or more legal doctrines have “overlapping and reinforcing coverage.” The reinforcing nature of this coverage means that, at least in some circumstances, the doctrines use substantially shared facts and concerns to generate shared outcomes that each individual doctrine could suffice to generate by itself. For example,

11 Golden, supra note 6, at 636.
12 Id.
in contract law, the duty of good faith and fair dealing in contract performance,\(^\text{13}\) the pre-existing duty rule,\(^\text{14}\) and doctrines of duress\(^\text{15}\) and unconscionability\(^\text{16}\) are redundant in the sense that each can act as independent grounds for finding unenforceable a contract modification extracted unfairly and without new consideration.\(^\text{17}\)

Importantly, doctrinal redundancy need not be complete. Legal doctrines can be partially redundant in the sense that their areas of coverage, use of facts, or outputs only partially coincide.\(^\text{18}\) Accordingly, the duty of good faith and fair dealing and doctrines of duress and unconscionability can have an operative effect in situations in which there is no pre-existing duty, and the pre-existing duty rule—however controversially—can prevent enforcement of promises in situations in which there is no bad faith, improper threat justifying a finding of duress or unconscionable conduct.\(^\text{19}\)

Indeed, partial redundancy is likely to be much more common and, policy-wise, more interesting than complete redundancy. When two legal doctrines are recognizably equivalent in all respects, their application, if not their very existence, will likely collapse into a single, unified analysis or a trivially wasteful and purely repetitive one. For example, in 2006, the U.S. Supreme Court enabled both of these potential effects by adopting a four-pronged test for permanent injunctions with two prongs—(1) a requirement for irreparable injury and (2) a requirement for inadequacy of legal

\(^{13}\) \textit{Restatement (Second) of Contracts} § 205 (Am. Law Inst. 1981) (“Every contract imposes upon each party a duty of good faith and fair dealing in its performance and its enforcement.”).

\(^{14}\) See \textit{id.} § 73 (“Performance of a legal duty owed to a promisor which is neither doubtful nor the subject of honest dispute is not consideration.”): \textit{cf. id.} § 89 (“A promise modifying a duty under a contract not fully performed on either side is binding (a) if the modification is fair and equitable in view of circumstances not anticipated by the parties when the contract was made.”).

\(^{15}\) See \textit{id.} § 175(1) (“If a party’s manifestation of assent is induced by an improper threat by the other party that leaves the victim no reasonable alternative, the contract is voidable by the victim.”).

\(^{16}\) See \textit{id.} § 208 (“If a contract or term thereof is unconscionable at the time the contract is made[,] a court may refuse to enforce the contract.”).

\(^{17}\) \textit{Cf. E. Allan Farnsworth, Contracts} § 4.22, at 272–73 (4th ed. 2004) (noting that the Uniform Commercial Code’s requirement of good faith and fair dealing imposes “the same standard as that suggested earlier under the liberalized rules on duress” and that “[t]his expanded concept of duress would make promises voidable in precisely those situations in which the pre-existing duty rule can be justified”).

\(^{18}\) Golden, \textit{supra} note 6, at 637–41. Further, there can be a phenomenon of probabilistic partial redundancy in accordance with which (1) it might be unclear whether two separately articulated legal doctrines are completely or only partially redundant and (2) this relationship might change as understandings of the doctrines and their contexts evolve. \textit{Id.} at 641.

\(^{19}\) See \textit{Farnsworth, supra} note 17, § 4.21, at 270 (observing that the pre-existing duty rule “does not . . . distinguish between the situation in which the contractor’s demand for more money is motivated merely by opportunism and greed and the situation in which the demand is prompted by discovery of circumstances or the occurrence of events that makes the contractor’s performance much more burdensome”).
remedies—that appear properly understood to be “one and the same.”

Unsurprisingly, trial courts have generally found that these two prongs rise or fall together, and it is unclear that their separation accomplishes much in one direction or another.

In contrast, partial redundancy can be more interesting in the sense that it more likely offers not only substantial benefits but also substantial costs. Costs tend to be greater than for complete redundancy because courts or other decision-makers must engage in somewhat differential analysis of each of two only partially redundant doctrines. On the other hand, benefits can also be greater because the lack of complete redundancy can mean that performance of a second, distinct analysis involving similar facts or concerns is more likely to add significant value than a second, completely redundant analysis would. In short, whereas complete redundancy between legal doctrines is commonly trivial in the sense of being substantially pointless and largely inconsequential, partial redundancy can frequently accomplish much, but details and context will determine whether, on net, what is accomplished is for good or ill.

B. Arguments for Redundancy and Anti-Redundancy

This Section recounts some of the major benefits and costs that can be associated with partial redundancy. These benefits and costs can be grouped in three basic areas: (1) effects of redundancy on the occurrence of legal “error,” which we take to mean a legal outcome that is socially undesirable; (2) redundancy’s association with efficiency in the development or administration of the law; and (3) positive or negative implications of redundancy for the flexibility and evolutionary potential of legal doctrine.

On the positive side, redundancy can add value even with respect to the efficiency of lawmaking and law application, despite common characterizations to the contrary. More specifically, when getting the “right” answer on any one of multiple partially redundant doctrines will lead to the “right” social outcome, redundancy can generate error-avoidance gains by increasing the probability that a court or other decisionmaker will arrive at the “right” result. Redundancy can be especially important in this regard when the primary social interest at stake has substantial priority over others, or when there is ambiguity or another significant source of error in applying

---

20 Golden, supra note 6, at 638–39 (quoting Mark P. Gergen et al., The Supreme Court’s Accidental Revolution? The Test for Permanent Injunctions, 112 COLUM. L. REV. 203, 209 (2012)).


22 Section I.B discusses reasons why partial redundancy can have such an advantage relative to complete redundancy.
one or more of the partially redundant doctrines.\textsuperscript{23}

In relation to law development or administration, redundancy can be associated with greater efficiency in at least two sets of circumstances. The first set of circumstances involves situations in which it is easier to develop a set of partially redundant legal doctrines to protect an identified set of interests than to develop a more streamlined or non-overlapping set of doctrines to achieve the same ends. The second set of circumstances involves situations in which deployment of a set of partially redundant doctrines makes a larger number of cases either easy to resolve or resolvable substantially earlier.\textsuperscript{24} Such earlier or easier resolution can occur because, under certain circumstances, one doctrine or another facilitates resolution of a legal dispute at an earlier stage than would be allowed by the others or because, even at a single decision point, a decisionmaker can prioritize application of the doctrine that offers the simplest path to a decisive legal conclusion.\textsuperscript{25}

Finally, redundancy can generate evolutionary and flexibility gains when a combination of different but partially redundant doctrines offers distinct opportunities for growth and adaptation of the law.\textsuperscript{26} Such opportunities can reflect either the fact that one doctrine’s securing of core legal interests makes more tolerable the relative fluidity of another doctrine, or the fact that each doctrine contributes its own “possibility frontier”—its own margins of ambiguity or uncertainty—along which courts or other legal decision-makers can effectively develop previously unsettled law.\textsuperscript{27}

As noted above, however, redundancy can have disadvantages along all of these dimensions as well. Under appropriate circumstances, these disadvantages justify the anti-redundancy rhetoric often featured in both policy debates and judicial decision-making.\textsuperscript{28} To a substantial extent, the potential disadvantages of redundancy mirror its potential benefits. First, redundancy can lead to increased errors when partially redundant doctrines

\textsuperscript{23} For a discussion of how the partial redundancy of improper appropriation and fair use helps safeguard free speech interests, see infra notes 200–208 and accompanying text.

\textsuperscript{24} For a discussion of improper appropriation as superior to fair use for early resolution of an important subset of cases, see infra notes 177–97 and accompanying text.

\textsuperscript{25} With a different conceptual framework of “hierarchy and specialization,” Henry Smith suggests similar advantages of retaining equity as a backstop to law, with law commonly featuring “simpler, more general and easier to follow” rules while “equity saves law from bad results in extreme situations.” Henry E. Smith, Equitable Defences as Meta-Law, in DEFENCES IN EQUITY 17, 25 (Paul S. Davies et al. eds., 2018).

\textsuperscript{26} For a discussion of how fair use can offer a safety net for adaptation of the law to new technological and economic developments, backing up the core function of the rule excluding protection for functional subject matter, see infra notes 132–36 and accompanying text.

\textsuperscript{27} Golden, supra note 6, at 704–07.

\textsuperscript{28} See id. at 631 (“Judges frequently cite anti-redundancy principles in interpreting legal documents, opining on the structure of legal doctrine or objecting to ‘relitigation’ of issues under ostensibly different legal headings.”).
erect successive barriers to achievement of a “right” outcome or when redundancy generates error-producing confusion about the proper meaning of partially redundant doctrines. Second, redundancy can increase the costs of law development or administration by straining attention or demanding greater attention to navigate, without error, the sometimes subtle or even obscure distinctions between doctrines that can be associated with partial redundancy. Moreover, redundancy can sometimes diffuse developmental efforts across a swath of legal doctrines, thereby preventing a potentially more efficient focus of legal decisionmakers’ attention on a smaller set of well-traveled doctrinal channels. Third, redundancy can generate evolutionary and flexibility losses, for example, by entrenching a set of legal outcomes that turn out to be increasingly undesirable with time.

In short, depending on the circumstances, redundancy can be a design feature or a design flaw. When smartly deployed, redundancy can add much value, but poorly designed or administered redundancy can yield the various ills highlighted by anti-redundancy rhetoric. Part II will explore how doctrinal redundancy plays out in a number of prominent circumstances in copyright law.

Before proceeding with Part II’s contextual analysis, however, we should note that one hurdle to proper analysis of the pluses and minuses of redundancy can come from the mis-categorization of various situations as involving complete redundancy or no redundancy whatsoever. Spurious assertions of redundancy perhaps most often come in the form of allegations of complete redundancy when there is, in fact, only partial redundancy. A classic textual example is the statement “war is war,” which had more than tautological meaning when penned by General Sherman during the month that his troops took Atlanta. The contrapositive phenomenon is a denial of even partial redundancy when there is a lack of complete redundancy. The

---

29 Id. at 704–05. For a discussion of evidence of judicial confusion about the relationship between the partially redundant rights of reproduction and of preparing derivative works, see infra text accompanying notes 301–93.
30 See Golden, supra note 6, at 668, 705 (explaining that “avoiding redundancy can be costly as well, perhaps particularly at the stage of generating and designing legal documents, doctrines, processes, or institutions”).
31 See id. at 672–73 (discussing the inverse in how anti-redundancy could “foster greater predictability and perhaps even accuracy in legal judgments”).
32 See id. at 706–07 (explaining redundancy’s evolutionary value).
33 See id. at 644 (defining spurious redundancy and explaining that often, complete redundancy is only partial).
34 See id. (invoking the example of “War is war” as a “classically redundant statement” (citing 2 William T. Sherman, MEMOIRS OF GENERAL WILLIAM T. SHERMAN 111 (1875) (quoting the end of a letter to General Halleck, which said, “If the people raise a howl against my barbarity and cruelty, I will answer that war is war, and not popularity-seeking”))).
35 See DANIEL SOLOW, HOW TO READ AND DO PROOFS: AN INTRODUCTION TO MATHEMATICAL THOUGHT PROCESS 30
fact that appellate judicial review is generally distinct from trial proceedings and commonly accepted as well-justified does not mean that aspects of that review are not at least partially redundant with the work of the trial court. In general, partial redundancy as we define it is likely to be much more prevalent than acknowledged by many discussions of redundancy. It is crucial to our perspective that redundancy be understood as a potentially multi-dimensional phenomenon—generally a question of degree, rather than a binary question of zero or one.

II. DOCTORAL REDUNDANCY IN COPYRIGHT

Copyright law is rich in partial redundancy. More specifically, copyright law features multiple sets of doctrines that address the same or overlapping concerns in various circumstances or otherwise perform overlapping functions. This Article explores four such sets: (1) the fair use doctrine and the separately articulated prohibition of copyrighting functionality; (2) the fair use doctrine and separately articulated limitations on what is considered improper appropriation copying; (3) limitations on copyrightable subject matter and copyright’s requirement of originality; and (4) copyright’s exclusive rights of reproduction and of making derivative works.

A. Fair Use and Functionality

Let us start by exploring the relationship between the rule that denies copyright protection to functional subject matter and the fair use doctrine, which excepts certain uses of copyrighted material from liability. The former rule is usually traced back to the 1879 decision, Baker v. Selden, in which the Supreme Court refused to extend copyright protection to a chart necessary for the use of a new accounting method. The decision stands for two main propositions.

The first proposition indicated by Baker is the perhaps common sense one that copyright protection cannot extend to functional or utilitarian subject matter. The Court’s example in Baker was a book describing the design of a machine. While the expression in the book is entitled to protection as a literary work, the mechanical design it describes is not

(1982) (observing that “NOT B implies NOT A” is the contrapositive of “A implies B” and that “the statement “NOT B implies NOT A” is true under the same conditions as “A implies B” (internal quotation marks omitted)).

36 See id. at 640–41 (illustrating partial redundancy through the example of trial and appellate courts).


38 Id. at 102; see Christopher Buccafusco & Mark A. Lemley, Functionality Screens, 103 Va. L. Rev. 1293, 1319 (2017) (describing Baker as being based on the principle that functional elements of expressive works do not receive protection).

39 Baker, 101 U.S. at 102.
If at all, the mechanical design may receive patent protection, subject to meeting the various substantive and procedural patentability requirements. Thus, one who copies the book may infringe upon its copyright, but one who reproduces the mechanical design described in the book does not.

The second *Baker* proposition, the so-called merger doctrine, raises the stakes. It instructs that, when expressive subject matter and functional subject matter are entwined, copyright protection must be denied to any expressive elements that are also part of, and hence for doctrinal purposes “merged” with, the relevant function. In *Baker*, for example, the disputed chart consisted of a combination of text and graphic elements that ordinarily would be classic copyrightable expression. But if a person wanted to use the new accounting system, she would apparently have had to use Selden’s chart or something very similar to it. In this sense, the expressive chart was the functional method or part thereof. Under such circumstances, the merger doctrine insists that no copyright protection subsists in the merged expressive elements. One way of restating this principle is that, when expressive and functional elements are merged, the non-copyrightability of function preempts copyright protection for expression.

Section 102(b) of the Copyright Act of 1976 codifies *Baker v. Selden’s*...
exclusion of functional subject matter. This is a loose codification, however. Section 102(b) simply lists “procedure[s], process[es], system[s], [and] method[s] of operation” as among the kinds of subject matter to which copyright protection does not extend. This formulation is broadly understood as incorporating preexisting case law, including the merger doctrine, on the exclusion of functional matter from copyright, but also as leaving much room for the courts to develop the exclusion in new contexts.

The fair use doctrine, loosely codified in § 107 of the Copyright Act, is currently the most important mechanism for restricting copyright’s coverage. Section 107 defines behavior which otherwise violates the owner’s exclusive rights as “not an infringement of copyright” if found to be fair use. The section lists several examples of purposes of using copyrighted works that may make the use fair, but the construal of these examples as merely illustrative has greatly diminished their importance. Section 107 also directs courts to consider four statutory factors: “(1) the purpose and character of the use . . . ; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used . . . ; and (4) the effect of the use upon the potential market . . . .” The American fair use doctrine, often contrasted with its British fair dealing cousin, is famously open-ended. In prescribing a broad standard accompanied by four factors, themselves fashioned as capacious standards, the doctrine contemplates the courts as the senior partner in fashioning the exact content of the doctrine and applying it to new contexts.

To understand the intersection between § 102(b) and fair use, consider Oracle v. Google. In this complex case, Oracle asserted copyright

47 Id.
48 See H.R. REP. No. 94-1476, at 57 (1976), as reprinted in 1976 U.S.C.C.A.N. 5659, 5670 (stating that the purpose of § 102(b) is to “restate, in the context of the new single Federal system of copyright, that the basic dichotomy between expression and idea remains unchanged”); Pamela Samuelson, Why Copyright Law Excludes Systems and Processes from the Scope of Its Protection, 85 TEX. L. REV. 1921, 1952 (2007) (analyzing portions of the House and Senate Reports relating to § 102(b)).
50 Id.
51 Id.
52 See, e.g., Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 561 (1985) (“This listing was not intended to be exhaustive . . . . The fact that an article arguably is ‘news’ and therefore a productive use is simply one factor in a fair use analysis.”); William F. Patry, PATRY ON COPYRIGHT §10.12 (2018) (“The preamble to Section 107 begins by enumerating illustrative methods of reproduction . . . and then enumerates illustrative purposes to which such reproductions are put.”).
protection in what computer professionals refer to as “libraries” of Application Programming Interfaces (“APIs”). API libraries can be described as compilations of preprogrammed packets of computer code and associated declarations that are used to “call” the implementing code, triggering the implementing code to perform specified functions. For those familiar with Microsoft Excel spreadsheets, it might help to think of an API declaration as a statement analogous to “=average(D5:D7),” a statement that instructs a computer to run code in the Excel program that will yield the average of entries in rows five through seven of column D of an Excel spreadsheet.

The specific background for Oracle’s complaint is as follows. Oracle’s predecessor, Sun Microsystems, developed the Java computer language. It also created packages of libraries for functions. Each function in these libraries is a pre-written segment of code that carries out a specific task—for example, a mathematical calculation, communication with a device, or a graphic operation. Rather than rewrite code, Java programmers can use these pre-fabricated functions to perform the relevant task. Oracle did not assert copyright in the Java language, and Google never copied the code in the libraries except for a very short, nine-line segment of code, which Google copied along with eight computer files. Instead, Google independently wrote its own libraries and made them available to programmers who wished to use the Java language to write software programs able to run on the Android mobile operating system. With respect to these libraries, which are the focus of our concern here, what Google did copy and Oracle argued was covered by copyright were the Java libraries’ APIs, namely: (1) the names and syntax for the declarations used to call

---

57 Id. at 1347, 1349.
58 Id. at 1349–50.
59 Id. at 1348 n.1.
60 Id. at 1348–49.
61 Id.
62 Oracle Am., Inc. v. Google Inc., 872 F. Supp. 2d 974, 978 (N.D. Cal. 2012) (“Google decided to use the Java language to design its own virtual machine via its own software and to write its own implementations for the functions in the Java API that were key to mobile devices.”), aff’d in part and rev’d in part, 750 F.3d 1339 (Fed. Cir. 2014).
63 Oracle, 750 F.3d at 1351 (“It is undisputed, however, that Google wrote its own implementing code, except with respect to: (1) the rangeCheck function, which consisted of nine lines of code; and (2) eight decompiled security files.”). The district court issued a judgment holding in part that Google had infringed the copyright for the eight copied files and the short segment of “rangeCheck” code, and the Federal Circuit affirmed these portions of the district court’s judgment. See Oracle, 750 F.3d at 1347–48 (affirming the holding of copyright infringement for the eight decompiled files and denying Google’s cross-appeal on the rangeCheck function). For Google’s infringement of the rangeCheck code, the district court had “entered judgment in the amount of zero dollars per the parties’ stipulation.” Oracle Am., Inc. v. Google Inc., No. C:10-3561, 2012 WL 9028839 (N.D. Cal. June 20, 2012).
64 Oracle, 750 F.3d at 1350–51 (describing Google’s development and release of its libraries as part of the Android platform).
associated functions, and (2) the relationships between the individual functions and their taxonomic structure.

In other words, Oracle argued that it had created a work subject to copyright protection that comprised not only the API declarations’ names and syntax, but also the complex and abstract informational structure embodied in the way that the numerous functions were organized, classified, and related to each other. The reason for Google’s copying of these elements is obvious: it wanted to swiftly mobilize the substantial community of Java programmers to produce Android software without having to learn a complex set of new conventions.

Oracle argued that the copied Java libraries’ APIs are protectable expression. One of Google’s main counterarguments was that the structure of the APIs is an unprotectable “command structure” under § 102(b), and that any expression embodied in them merged with their functional aspect.

At the heart of this highly technical dispute was a battle over interoperability. Interoperability is the ability of technology—hardware or software—to work or communicate with other technology. Especially in the context of computer software, one of the important functions of § 102(b)’s preclusion of copyrighting functionality is to ensure access to interoperability. Just as a highly sophisticated new cellphone that cannot be connected to existing networks might as well be used as a doorstop, even an otherwise-superior spreadsheet application will be seriously disadvantaged if it cannot read the file format used by other established

---

65 Id. at 1351. The court’s opinion confusingly refers to the syntax of the libraries’ declarations as “the declaring code.” Id. The declarations are not the code of the libraries, however, but simply the conventions used in calling them. For example, the declaration “public static int max(int x, int y)” is used to call a function that has certain traits and returns an integer number that is the larger of two input integers. See id. at 1349–50 (explaining declarations and how functions are “called”).

66 Id. at 1351. The opinion refers to this aspect as the “structure, sequence, and organization” (SSO) of the libraries. Id.

67 Oracle, 872 F. Supp. 2d at 978 (“As to the 37 packages at issue, Google believed Java application programmers would want to find the same 37 sets of functionalities in the new Android system callable by the same names as used in Java.”).

68 Oracle, 750 F.3d at 1353.

69 Id. at 1353, 1356.

70 A standard definition of interoperability in the context of information and communication technology is “[t]he capability to communicate, execute[] programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units.” Int’l Electrotechnical Comm. [IEC], Information Technology Vocabulary, at 6, ISO/IEC 2382-1 (3d ed. 1993); see also Pamela Samuelson, Are Patents on Interfaces Impeding Interoperability?, 93 MINN. L. REV. 1943, 1946–48 (2009) (defining interoperability).

71 See Lexmark Int’l, Inc. v. Static Control Components, Inc., 387 F.3d 522, 536 (6th Cir. 2004) (“[L]ock-out’ codes fall on the functional-idea rather than the original-expression side of the copyright line.”); Sega Enters. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1522 (9th Cir. 1992) (“[F]unctional requirements for compatibility . . . are not protected by copyright.”); see also Samuelson, supra note 48, at 1976 (advocating the conclusion that “systematic assemblages of information such as specifications of interfaces necessary to achieve interoperability are unprotectable under § 102(b)”).

applications to store data, and thus requires re-entry of the data as a prerequisite to manipulation by the application. Standard economic reasoning provides two related explanations why interoperability is an important policy goal: network externalities and lock-in effects due to sunk investment.

Positive network externalities are the beneficial effects enjoyed by each user of a resource as a result of use by others.\(^2\) The classic example is a telephone network where the value of a device is virtually zero if operated in isolation but steadily increases in relation to the number of users connected to the network.\(^3\) When a particular resource, such as a well-established network or standard, is subject to high network effects, exclusion of newcomers from enjoying these externalities is likely to erect substantial barriers to entry.\(^4\) Consider the plight of an entrant who tries to establish a new local telephone network but is denied interconnectivity with the national network. How many customers will sign up for the entrant’s service?

Lock-in effects often work in conjunction with network externalities. At the heart of this phenomenon are sunk investments that generate path dependence\(^5\) or what scientists or engineers might describe more specifically as a form of hysteresis.\(^6\) Consider a firm that invests heavily in training its employees in mastering the protocols for operating a machine central for its business. Sometime later, a new machine superior in performance enters the market. If the new machine could be substituted seamlessly for the old one without requiring new employee training, the new machine’s acquisition and deployment would raise the firm’s profit margin


\(^{3}\) Katz & Shapiro, supra note 72, at 424.

\(^{4}\) See Mark A. Lemley & David McGowan, Legal Implications of Network Economic Effects, 86 CAL. L. REV. 479, 523 (1998) (“To the extent intellectual property rights confer ownership interests in a strong network standard, they may create durable market power in network markets . . . . But if one firm owned the rights to produce keyboards with a QWERTY layout, it could exclude others from obtaining the benefits of the standard.”).


\(^{6}\) See, e.g., C.J. Adkins, Equilibrium Thermodynamics 10 (3d ed. 1983) (“Friction is a common cause of hysteresis.”); Neil W. Ashcroft & N. David Mermin, Solid State Physics 722 (1976) (defining “hysteresis” as how, after a piece of iron has been magnetized through application of a magnetic field, it can “become[] necessary to apply a rather strong field in the opposite direction to restore the unmagnetized configuration”).
from \( P_1 \) to \( P_2 \), where \( P_2 > P_1 \) despite the costs of such acquisition and deployment. Suppose, however, that the new machine requires expensive new training costing the amount \( T \) in order for the machine to be used successfully as a substitute for the earlier machine. Expected profit from using the new machine will then be only \( P_2 - T \). If \( P_2 - T \) is less than \( P_1 \), a rational profit-maximizing firm will not adopt the new machine: the firm will be “locked into” the inferior technology. Significantly, such “lock-in” can occur even when, in the absence of the need for the new investment in training—for example, because existing operating protocols can be used without alteration for the new machine—both the firm’s private interest and the social welfare calculus point in favor of adopting the newer technology.

Network and lock-in effects due to sunk investment have a common adverse effect: they create barriers to entry.\(^77\) If significant enough, they can inefficiently chill competition and the introduction of socially beneficial innovations.\(^78\) Various features of the computer software industry—including the highly cumulative nature of its technology, the industry’s reliance on networks of users, operators, and developers, and associated reliance on standards—make the industry an especially fertile ground for network externalities and lock-in effects.\(^79\) And by allowing new software to use or interact with key aspects of existing software, interoperability can often dissolve or mitigate associated entry barriers for new technologies or new industry players.\(^80\) Copyright law uses §102(b) and the related merger doctrine as mechanisms for preventing these pernicious effects in various expressive fields and in software in particular.\(^81\) By designating information structures or expressive elements needed for interoperability as functional and therefore unprotected, §102(b) and the merger doctrine ensure access

\(^{77}\) See Lemley & McGowan, supra note 74, at 522–23 (“[C]ompetition between different networks may be counterproductive in certain circumstances.”).

\(^{78}\) See id. at 522 (“[N]etwork effects may foreclose competition entirely or limit effective competition to that occurring between members of the same network.”).


\(^{80}\) See Peter S. Menell, Rise of the API Copyright Dead?: An Updated Epitaph for Copyright Protection of Network and Functional Features of Computer Software, 31 HARV. J.L. & TECH. 305, 466 (2018) (arguing that copyright should allow “freedom to use functional features and to develop interoperable products” while patent law should protect functional elements of software to ensure proper incentive for their development and avoid certain negative effects related to network externalities).

\(^{81}\) Cf. Atari Games Corp. v. Nintendo of Am. Inc., 975 F.2d 832, 839–40 (Fed. Cir. 1992) (observing that although computer code is generally copyrightable expression of a way to implement a process, “[i]f the patentable process is embodied inextricably in the line-by-line instructions of the computer program, . . . then the process merges with the expression and precludes copyright protection”).
to network externalities and help prevent inefficient lock-in.\textsuperscript{82}

Notably, however, the interoperability interest in \textit{Oracle} was of a peculiar kind.\textsuperscript{83} The Federal Circuit decided the case on the assumption that Google’s libraries did not make preexisting Java code interoperable with the Android operating system.\textsuperscript{84} Indeed, Oracle complained that Google was compromising the “write once, run anywhere” hallmark of Java by failing to provide for such interoperability.\textsuperscript{85} If Google could not claim that the copying was necessary to achieve interoperability between the Android system and preexisting Java code, then what interoperability interest was it asserting?

The remaining argument—that relevant interoperability was at stake—focused on Google’s interest in drawing on the significant resource of the existing Java programming community in order to ensure fast and robust application development for the Android environment.\textsuperscript{86} This could be called an interest in the interoperability of Android with preexisting programming potential. There remains the question of why the use of preexisting APIs should be viewed as “functional.” After all, Google could have used new and different declaration syntax and a different library taxonomy in providing the desired prefabricated functionalities to Java programmers. How then are the syntax and taxonomy of APIs functional?

The answer lies in the fact that the functional purpose of an API is to “interface” with programmers, facilitating their development of application programs.\textsuperscript{87} In this context, demanding that Google use a different syntax and a different taxonomy is analogous to telling a developer of a new cellphone both to create its own network and to induce a critical mass of users to switch to the new network. To prevent programmers from being locked into traditional Java by virtue of their sunk investments in learning the syntax and structure of its libraries, Google needed more than perfectly workable libraries. It needed its libraries to make programmers’ sunk investment transferable. And to achieve that, it needed to follow established conventions from the Java libraries. Anyone who wanted effective interoperability with established programming capital needed to use Oracle’s associated expression, just as anyone who wanted to use Selden’s accounting method needed to use his chart.

Whether the peculiar kind of interoperability asserted by Google should

\textsuperscript{82} See Lotus Dev. Corp. v. Borland Int’l, Inc., 49 F.3d 807, 819–20 (1st Cir. 1995) (Boudin, J., concurring) (discussing § 102(b) as one doctrinal alternative for addressing issues of lock-in and network externalities in computer software).

\textsuperscript{83} Oracle Am., Inc. v. Google Inc., 750 F.3d 1339, 1353 (Fed. Cir. 2014).

\textsuperscript{84} Id. at 1371.

\textsuperscript{85} Id. at 1350.

\textsuperscript{86} Id. at 1371 (“Google wanted to capitalize on the fact that software developers were already trained and experienced in using the Java API packages.”).

\textsuperscript{87} Id. at 1349.
be safeguarded from copyright protection is open for debate (although we believe the answer is “Yes”).\(^88\) What matters for our purposes is the Court’s analysis of the interoperability argument and how this analysis reflected an understanding of the relationship between § 102(b) and fair use.

The Federal Circuit rejected Google’s argument by developing a vision of a strict separation between § 102(b) and fair use.\(^89\) Section 102(b), the court said, is about the eligibility of certain subject matter created by authors for copyright, while fair use is about whether certain actions by users are infringing.\(^90\) As such, the two never intersect in a theoretical sense, and each is applied from a very different perspective. Section 102(b) is applied ex ante from the point of view of the author of the protected work at the moment of creation.\(^91\) Its analysis asks, whether at that time and from the perspective of the author, a particular expressive form was functional or merged with functional elements.\(^92\) Fair use, and fair use alone, is applied ex post from the perspective of a user of a copyrighted work to examine whether at that point there exist any circumstances that excuse the copying.\(^93\)

This framing of the relationship between § 102(b) and fair use sealed the fate of Google’s § 102(b) argument by requiring that this argument be assessed from an ex ante perspective. When Sun created the Java libraries, it was free to choose any syntax and structure for them, unhampered by any

\(^88\) Many commentators support allowing copying for purposes of interoperability in circumstances such as those in Oracle. See, e.g., Pamela Samuelson, Three Fundamental Flaws in CAFC’s Oracle v. Google Decision, 37 EUR. INTELL. PROP. REV. 702, 702 (2015) (arguing, in part, that the Oracle court misinterpreted aspects of § 102); Peter S. Menell, API Copyrightability Bleak House: Unraveling and Repairing the Oracle v. Google Jurisdictional Mess, 31 BERK. TECH. L.J. 1515, 1566 (2016) (discussing Ninth Circuit decisions that held that “the code necessary for interoperability is uncopyrightable”); Wendy J. Gordon, How Oracle Erred: The Use/Explanation Distinction and the Future of Computer Copyright, in COPYRIGHT LAW IN AN AGE OF LIMITATIONS AND EXCEPTIONS 375, 375 (Ruth L. Okediji ed., 2017) (arguing that the “Oracle court may have been asking the wrong question” when determining copyrightability).

\(^89\) Oracle, 750 F.3d at 1354. The Federal Circuit’s driving of a strong doctrinal wedge between § 102(b) and fair use analysis resonates with similar clarification-and-compartmentalization moves that have marked the historical development of patent law, an area of U.S. law for which the Federal Circuit has virtually exclusive authority to hear initial appeals. See Golden, supra note 6, at 684–90 (describing a long-term trend toward doctrinal compartmentalization in patent law). But, whatever the value of the Federal Circuit’s helping to continue this long-term trend in patent law, one can wonder about the wisdom of the court’s transferring compartmentalization tendencies to copyright, a field in which the court is less expert and can expect generally to be applying the law of various regional circuits. See Oracle, 750 F.3d at 1353 (noting that “[c]opyright issues are not exclusively assigned to the Federal Circuit” and that, accordingly, “[t]he parties agree that Ninth Circuit law applies”).

\(^90\) Oracle, 750 F.3d at 1372 (The ability of a copier to achieve interoperability “might be relevant to the fair use inquiry,” but it is “irrelevant to the copyrightability” analysis.).

\(^91\) Id. at 1361.

\(^92\) Id. (“[C]opyrightability and the scope of protectable activity are to be evaluated at the time of creation, not at the time of infringement.”).

\(^93\) Id. at 1369 (dismissing the relevance for § 102(b) analysis of previous decisions that recognized a copier’s interest in interoperability because they were “focused on fair use”).
functionality, interoperability, or merger concerns.\footnote{Id. at 1356.} At the moment of their creation, there was no preexisting network, either technological or human, that necessitated a particular structure for the Java libraries to achieve effective interconnectivity with an already developed community of Java programmers.\footnote{Id. at 1371.}

Does the Federal Circuit’s neat and tidy vision of strict separation between functionality and fair use work? It does work perfectly, but only in von Jhering’s Heaven of Legal Concepts.\footnote{Rudolf von Jhering, Im juristischen Begriffshimmel, in SCHIERZ UND ERNST IN DER JURISPRUDENZ (1884) (Ger.), translated in In the Heaven for Legal Concepts: A Fantasy, 58 TEMP. L.Q. 799, 799 (1985). Felix Cohen borrowed von Jhering’s metaphor and made it a common designation in American legal thought of overly conceptuast legal thinking disconnected from considerations of purpose and policy. See Felix Cohen, Transcendental Nonsense and the Functional Approach, 35 COLUM. L. REV. 809, 809 (1935) (explaining that the heaven of legal concepts in von Jehring’s work is a place where legal concepts exist in their “absolute purity, freed from all entangling alliances with human life”).} The Federal Circuit’s reasoning is unburdened by any attempt to understand the underlying purpose of § 102(b) or the socioeconomic and technological context in which it operates. But establishing the proper understanding of § 102(b) and its relationship to fair use can only be done in light of these doctrines’ relevant purposes.

The denial of copyrightability to functional subject matter and the associated merger doctrine are about the freedom of subsequent users to use such subject matter (unless it is patented).\footnote{Lexmark Int’l, Inc. v. Static Control Components, Inc., 387 F.3d 522, 535 (observing that when expression and ideas merge, “copyright protection does not exist because granting protection to the expressive component of the work necessarily would extend protection to the work’s uncopyrightable” elements).} What matters here is not some metaphysical inquiry about whether the material in which copyright is asserted is inherently functional or expressive. The relevant inquiry is a pragmatic one: whether copyright in the relevant elements amounts to exclusive control of actions—which as a matter of social practice are within the realm of functional utility—rather than just the sort of creative expression that is copyright’s proper domain. When the lens is switched from metaphysics to pragmatism, the inquiry can hardly be undertaken without considering the effects on users. This is doubly true when the pertinent concern of “functionality” is interoperability. As explained, the policy concern here is about network externalities and lock-in effects in the interface between human behavior and technological networks. These

---

94 Id. at 1356.
95 Id. at 1371. The Federal Circuit went on to analyze the framework for a fair use analysis of the case and remanded it to the district court for further proceedings on this issue. Id. at 1373–77. In March 2018, the Federal Circuit reversed the district court on this issue and found Google’s actions not to constitute fair use. Oracle Am., Inc. v. Google LLC, 886 F.3d 1179, 1211 (Fed. Cir. 2018). Google has filed a petition for reharing en banc, which is currently pending. Petition for Rehearing En Banc, Oracle Am., Inc. v. Google LLC, Nos. 17-1118, -1202 (Fed. Cir. petition dated May 29, 2018), available at https://patentlyo.com/media/2018/05/Oracle-Am.-v.-Google-LLC-Rehearing-Petition.pdf.
97 See Lexmark Int’l, Inc. v. Static Control Components, Inc., 387 F.3d 522, 535 (observing that when expression and ideas merge, “copyright protection does not exist because granting protection to the expressive component of the work necessarily would extend protection to the work’s uncopyrightable” elements).
effects can shield an incumbent from competition and create a barrier for the development and entry of innovative products. Network and lock-in effects, however, almost always develop over time. There was no network effect when, in 1876, Bell first spoke the words “Watson come here” over a telephone line, and any lock-in effect was of trivial social magnitude just after the first typist learned to use a QWERTY keyset on a typewriter.

There is nothing unusual about considering ex post effects on users as part of IP subject matter rules. The genericity doctrine in trademark precludes protection for marks that designate for the public not a specific source of goods, but a whole type of goods. The doctrine of subsequent genericity is founded on the idea that a particular mark, such as “cellophane,” “thermos,” or “Murphy bed,” might initially designate only a particular source of goods, but over time end up being generic. At that point of subsequent genericity, trademark protection is lost.

Likewise, in copyright, one finds the scènes à faire doctrine. Under this rule, common or stock elements within a particular expressive genre do not receive copyright protection. Under this doctrine, even the first creator of a shoot-out at high noon scene in a western movie cannot control such

98 TIm Wu, The Master Switch: The Rise and Fall of Information Empires 22 (2010).
99 See Yun-chien Chang & Henry E. Smith, An Economic Analysis of Civil Versus Common Law Property, 88 Notre Dame L. Rev. 1, 18 (2012) (noting reference to “the QWERTY typewriter keyboard” as an example of lock-in, but also noting dispute over whether there was a substantially superior alternative on which users would have otherwise converged); see also David, supra note 76, at 335–36 (discussing the QWERTY “lock in” effect); Liebowitz & Margolis, supra note 76 (responding to David’s work regarding QWERTY history); Paul A. David, At Last, a Remedy for Chronic QWERTY-Skepticism! (Sept. 31, 1999) (unpublished manuscript) (on file with author) (responding to critiques of his work on QWERTY).
100 For a general argument that the passage of time affects the policies underlying copyright, see Joseph P. Liu, Copyright and Time: A Proposal, 101 Mich. L. Rev. 409, 411–12 (2002) (concluding that the strength and impact of theoretical justifications for copyright law are quite directly affected by the passage of time). In this article, however, Liu suggests adjusting copyright’s scope to the passage of time through the fair use doctrine, not subject matter rules. Id.
101 See Singer Mfg. Co. v. June Mfg. Co., 163 U.S. 169, 185 (1896) (explaining that when a patent expires, the public has a right to make the property in the same manner it was made during the patent); DuPont Cellophane Co. v. Waxed Prods. Co., 85 F.2d 75, 80 (2d Cir. 1936) (“The fact that it had registered ‘Cellophane’ as a trademark would give it no right to monopolize a term useful to designate a commercial article.”); 2 J. Thomas McCarthy, McCarthy on Trademarks and Unfair Competition § 12.1 (5th ed. 2018) (“[A] generic name of a product can never function as a trademark to indicate origin.”).
102 See, e.g., Murphy Door Bed Co. v. Interior Sleep Systems, Inc., 874 F.2d 95, 97 (2d Cir. 1989) (“We hold that ‘Murphy bed’ is a generic term, having been appropriated by the public to designate generally a type of bed.”); King-Seeley Thermos Co. v. Aladdin Indus., 321 F.2d 577, 579 (2d Cir. 1963) (noting that over a period of approximately fifty years, the term “thermos” became widely used to describe a generic type of insulation).
In Oracle, Google tried repackaging essentially the same interoperability argument as a scènes à faire claim. The court dismissed it for the same reasons it dismissed the § 102(b) claim—i.e., that the scènes à faire doctrine focuses “on the circumstances presented to the creator, not the copier.”

But how can scènes à faire be analyzed except from the ex post perspective of users? To be scènes à faire, an expressive element must be a stock element of a genre, and both a genre and what counts as stock elements within it almost inevitably develop over time. It is always fundamental early or earlier works that define what stock elements are from the point of view of subsequent creators. And to apply the doctrine from the ex ante point of view—looking from the moment where the earlier works were about to be created and gradually form the genre—would make the doctrine incoherent and defeat its purpose. Imagine that whether a train robbery scene is scènes à faire in western movies was analyzed from the point of view of Edwin S. Porter making his The Great Train Robbery in 1903. The same logic applies to § 102(b) in general and interoperability in particular. The collective dynamics of cultural and technological networks—especially those which are the concern of copyright policy—often develop over time. To restrict § 102(b) to the ex ante perspective of the creator of the protected work is to gut it.

Fine, the skeptic might say, Oracle may gut §102(b), but there is no need to worry because fair use can still effectively protect interoperability by immunizing subsequent users from liability when such a concern arises. To push the argument one step further: this arrangement may be preferable because it avoids the inefficient and confusing redundancy that would occur if choices facilitating interoperability were shielded by both fair use and § 102(b). Therefore, it is better to let fair use take care of users’ ex post concern and push functionality to a marginal or non-existent role on such an issue.

This anti-redundancy argument fails to account for how the relevant doctrines operate in practice. Fair use is a broad, open-ended standard. As a result, applying it often entails a complex and frequently information-intensive and expensive inquiry, one reaching far beyond the functionality,

---

104 See 4 NIMMER & NIMMER, supra note 103, § 13.03[B][4] (“[C]ertain patterns and situations are bound to recur.”).
105 750 F.3d 1339, 1363 (Fed. Cir. 2014).
106 Id. at 1364.
107 Swirsky v. Carey, 376 F.3d 841, 846 n.11 (9th Cir. 2004).
108 The Great Train Robbery is a 1903 silent film. THE GREAT TRAIN ROBBERY (Edwin S. Porter 1903). It is widely recognized as featuring ground breaking movie-making techniques, and many film historians consider it the first Western film. See NORMAN O. KEIM, OUR MOVIE HOUSES: A HISTORY OF FILM & CINEMATIC INNOVATION IN CENTRAL NEW YORK 17 (2008).
or non-functionality, of a particular form of expression.\textsuperscript{109} Because of this complexity and the open-ended nature of the doctrine, the fair use doctrine’s case-by-case outcomes can be hard to predict.\textsuperscript{110} This remains so despite the mitigating effect, pointed out by some scholars, of observable patterns in fair use decisions.\textsuperscript{111} By comparison with fair use inquiries, application of § 102(b) tends to be much more focused in terms of relevant facts and much more liable to development of relatively straightforward, rule-like precedent such as a potential rule that APIs with substantial network externalities are presumptively functional. Finally and importantly, fair use and § 102(b) ordinarily involve different assignments of the ultimate burden of proof: whereas argument on the basis of § 102(b) is generally understood to implicate preconditions for copyrightability on which copyright holder bears the burden,\textsuperscript{112} the Supreme Court has indicated fair use to be an affirmative defense for which the defendant bears the burden.\textsuperscript{113} Consequently, not only may a party facing a potential charge of copyright infringement predict with far greater assurance that it will prevail on the basis of functionality, rather than fair use, but that party also may expect to prevail on functionality


\textsuperscript{110} Pierre N. Leval, \textit{Toward a Fair Use Standard}, 103 HARV. L. REV. 1105, 1132 (1990) (“Whether [a] taking[ ] will pass the fair use test is difficult to predict. It depends on widely varying perceptions held by different judges.”).


\textsuperscript{112} See Samuelson, supra note 5, at 438 (“Courts generally put the burden of proof of non-merger on plaintiffs because to establish a prima facie case of infringement requires proof that the defendant appropriated original expression from the plaintiff’s work.”). The extent to which § 102(b) is applied as a precondition for copyrightability or as an affirmative defense is in fact a matter of some contention. In \textit{Oracle}, the Federal Circuit acknowledged some disagreement on this question in relation to application of the related doctrines of merger and scènes à faire. 750 F.3d at 1358. However, the particular case law of the U.S. Court of Appeals for the Ninth Circuit was understood to govern copyright issues in \textit{Oracle} and the Federal Circuit determined that the Ninth Circuit had “clear[ly]” held that these doctrines are affirmative defenses. \textit{Id.} at 1353, 1358. This approach received scholarly criticism for ignoring subtleties introduced by § 102(b)’s relevance, under previous courts’ approaches, to both threshold questions of copyrightability and an infringement analysis that distinguishes between copyrighted and non-copyrighted material. See Menell, supra note 89, at 1569 (criticizing the Federal Circuit for rigidly focusing on Ninth Circuit cases that treated the two doctrines as defenses to infringement rather than copyrightability doctrines).

\textsuperscript{113} Campbell v. Acuff–Rose Music, Inc., 510 U.S. 569, 590 (1994) (observing that “fair use is an affirmative defense”). One might question whether fair use should be an affirmative defense. Lydia Pallas Loren, \textit{Fair Use: An Affirmative Defense?}, 90 WASH. L. REV. 685, 688 (2015) (contending that the Supreme Court “should conclude that fair use is not an affirmative defense but is a mere defense”). We do not pursue that question as the key point for our overall concern with the structural relationships between legal doctrines, but simply that the existence of redundant inquiries such as those into fair use and compliance with § 102(b) enables the sort of differential assignment of burdens discussed in the text.
substantially earlier and more cheaply in any ensuing litigation—whether on a motion to dismiss based on pleadings or on summary judgment—rather than after a full trial.\textsuperscript{114}

The implications of the above are twofold. First, the prospect of mounting a fair use defense often offers parties who fear allegations of copyright infringement no more upfront than the right to engage in a long, expensive legal process with an uncertain result.\textsuperscript{115} The predictable result can be the “chilling” and likely overdeterrence of activity for which a strong fair use defense can be mounted.\textsuperscript{116} The Federal Circuit epilogue in \textit{Oracle} demonstrates this point.\textsuperscript{117} When the Federal Circuit rejected Google’s § 102(b) argument, the court left open the possibility of exemption for liability based on fair use.\textsuperscript{118} During the original trial, the jury had deadlocked on this issue.\textsuperscript{119} The Federal Circuit responded with guidance for the fair use analysis and remanded the case for further proceedings.\textsuperscript{120} Back in the trial court, the jury found Google’s use of the Java APIs to be fair, and the district court denied Oracle’s motion to overturn the verdict by a judgment as a matter of law.\textsuperscript{121} The district court judge supported this decision with detailed legal reasoning in favor of upholding the fair use outcome.\textsuperscript{122}
court, overturned the jury’s verdict, and found no fair use.\textsuperscript{123} In the course of reaching this endpoint, the Federal Circuit held that a conclusion that “Google’s use of the API packages [was] anything other than overwhelmingly commercial” lacked any “substantial evidentiary support,”\textsuperscript{124} that Google’s use was “not transformative as a matter of law,”\textsuperscript{125} that “no reasonable jury could conclude that what was copied was qualitatively insignificant,”\textsuperscript{126} and that “no reasonable jury could have concluded that there was no market harm to Oracle from Google’s copying.”\textsuperscript{127} The Federal Circuit effectively disagreed with the trial court on all four of these issues, two of which—the transformative nature of the use and market harm—are commonly considered the two most outcome-determinative elements of the fair use analysis.\textsuperscript{128}

What is most relevant for the current discussion, however, is not the Federal Circuit’s reasoning on the second appeal or this reasoning’s result. Instead, what is most relevant is the tortuous procedural route to this fair use result, a route that has involved two jury trials and two appeals, and after Google’s filing for a rehearing en banc,\textsuperscript{129} is still incomplete. Google has been able to follow this winding and expensive path to its so far bitter end. Other less wealthy or established players are more likely to quit earlier, rather than fight for so long to vindicate a fair use argument. Alternatively, anticipating a potential charge of copyright infringement and the need for a lengthy and costly legal battle to prevail against it, they might avoid any action that would risk such a charge in the first place.

This conclusion leads to the second point: the chilling effect of the complex and uncertain fair use doctrine does not affect all parties equally. Well-established and well-financed players are generally better placed to undertake an expensive and risky fair use defense, a course that will at least offer the relatively certain benefit of delaying any unfavorable final judgment.\textsuperscript{130} But smaller, less-established players are not as well positioned

\textsuperscript{123} Oracle, 886 F.3d at 1211.
\textsuperscript{124} Id. at 1198.
\textsuperscript{125} Id. at 1199.
\textsuperscript{126} Id. at 1207.
\textsuperscript{127} Id. at 1209. For an argument that a reasonable jury could have upheld Google’s fair use defense, see Pamela Samuelson & Clark D. Asay, Saving Software's Fair Use Future, 31 HARV. J.L. & TECH. 535, 557 (2018).
\textsuperscript{128} See Beebe, supra note 112, at 584 (“[T]he outcomes of factors one and four very strongly correlated with the test outcome and fairly strongly correlated with each other.”). More recent empirical scholarship has tended to emphasize the relative importance of the transformativeness inquiry under the first fair use factor in comparison to the market-harm inquiry under the fourth factor. See, e.g., Netanel, supra note 112, at 745 (explaining that today the most important factor is transformative use); Matthew Sag, Predicting Fair Use, 73 OHIO ST. L.J. 48, 76 (2012) (discussing the possibility that transformative use is the most important factor).
\textsuperscript{129} See supra note 94.
\textsuperscript{130} HANOCH DAGAN, PROPERTY: VALUES AND INSTITUTIONS 148 (2011) (arguing that broad standards tend to generate regressive results by favoring repeat players).
to survive the process and are more likely to forgo the defense altogether.\textsuperscript{131} This is particularly worrisome when much concern about interoperability revolves around maintaining robust entry to markets or, in other words, allowing dynamic upstarts to challenge an entrenched technological incumbent. Here, an unpredictable doctrine that is expensive to apply likely disfavors exactly those who need its protection the most. In contrast, by relying on a more streamlined inquiry with the burden on the copyright holder, robust enforcement of § 102(b) can provide such upstarts much more reassurance.

Section 102(b), then, has an important role to play despite its partial overlap with fair use, meaning that both may lead to similar results on the basis of similar circumstances. Perhaps, however, we should try the opposite tack and eliminate redundancy by taking fair use out of the business of helping to shield steps taken to achieve interoperability. Is the optimal arrangement a mirror image of the Federal Circuit’s version of strict separation in Oracle? Not quite. For reasons indicated above, § 102(b) should play the lead role here, but fair use still has an important function.

To begin with, a fair use doctrine that partially overlaps with § 102(b) analysis can provide a variety of safety-net functions. The doctrine can be an error-correcting failsafe in technology-intensive, complex cases such as Oracle, where juries or even judges might find it difficult to digest the intricacies of relevant technologies, or of the legal distinction between function and expression. In other words, in certain cases that should have been caught in the functionality net but escaped it, fair use can “save the day” by generating the right outcome—exemption from infringement liability—even when a court has erred in finding relevant subject matter to be copyrightable. Moreover, notwithstanding the generally more unwieldy nature of fair use analysis, there can be some circumstances—such as where there is an on-point, precedential finding of fair use—where the fair use doctrine might even provide an easier route to a conclusion of no liability.

More importantly, in yet another set of cases, fair use effectuates the freedom to achieve interoperability where other doctrinal levers, including § 102(b), currently cannot. Often, extracting unprotectable functional information necessary for interoperability requires some copying of protected expressive materials. In \textit{Sega Enterprises Ltd. v. Accolade, Inc.},\textsuperscript{132} for example, a video game developer wanted to independently create games compatible with the Sega console.\textsuperscript{133} To achieve that result, it needed access to the communication protocols of the console, information that is


\textsuperscript{132} 977 F.2d 1510 (9th Cir. 1992).

\textsuperscript{133} See id. at 1514–15 (explaining the developer’s attempts to make a compatible game).
unprotected by copyright but was not readily available. As a practical matter, the only way for the developer to obtain the unprotected communication protocols of the console was to reverse-engineer them by decompiling and examining copyrighted code of Sega games, a process that inevitably involved temporary copying of the games’ code. This game code was not functional under § 102(b) and hence was fully copyrightable. In principle, copyright law might have permitted temporary copying of such copyrighted code on grounds that it is not even a prima facie violation of the reproduction right when incidental to a purpose such as Accolade’s. But in the United States, that route was foreclosed long ago by courts and is unlikely to be reopened. In Sega, the court found a lack of copyright liability by deploying the fair use doctrine to exempt such ancillary copying for purposes of gaining access to functional information necessary for interoperability.

In this auxiliary role as a support for functionality rules under § 102(b), it is exactly the open-ended nature of fair use that is crucial. In a complex and fast-changing technological and economic environment, it is hard to predict the exact set of circumstances or strategies that might arise and frustrate the purpose of rules meant to protect important policy goals such as broadly available interoperability. The flexible fair use doctrine can enable the “right” social result in such cases where otherwise easier-to-use rules are too rigid to be extended, and legislative involvement, commonly plodding, is inadequate. What emerges, then, is the possibility of a particular kind of beneficial redundancy structure: a set of rules under § 102(b) that in a broad array of circumstances generates the “right” social result based on a relatively limited set of facts, along with a broad standard that acts as a failsafe for both the limitations of the § 102(b) rules and for errors in their application.

In short, fair use still has an important role to play alongside § 102(b). It

134 Id. at 1515 (discussing how Accolade engineers “experimented to discover the interface specifications for the Genesis console”).
135 Id. at 1520 (finding fair use because “disassembly [of copyrighted object code was] the only means of gaining access to the relevant unprotected aspects of the program.”).
136 Id. (“[T]he copyright in a computer program extends to the object code version of the program.”).
137 Another alternative is to have a specific statutory rule allowing copying for purposes of interoperability under such circumstances. This is the strategy adopted in the European Union. Council Directive 91/250/EEC of 14 May 1991 on the Legal Protection of Computer Programs, art. 6, 1991 O.J. (L 122) 0042.
138 Sega, 977 F.2d at 1519 (holding “intermediate copying” of code for purposes of interoperability to be infringing).
139 Id. at 1514 (noting that the copying of a copyrighted computer code to gain access to unprotectable functional elements is fair use when “the person seeking the understanding has a legitimate reason for doing so and when no other means of access to the unprotected elements exists”); see also Sony Comput. Entm’t v. Connectix Corp., 203 F.3d 596, 602–09 (2000) (finding the intermediate copying of BIOS code in the course of reverse engineering a video game console to be fair use).
can correct for errors, occasionally offer an easier route to an identical result, and provide auxiliary protection for social interests such as interoperability in the face of dynamic and varied circumstances. When informed by concern with underlying policies and implemented with sensitivity to how relevant doctrines function in practice, partial redundancy between § 102(b) limitations and the fair use doctrine is beneficial. Moreover, for purposes of exempting activities associated with securing a functional end such as interoperability, a structural combination in which § 102(b) is the primary mechanism and the fair use doctrine is a more general backstop emerges as a preferred design.

B. Fair Use and Improper Appropriation

To establish copyright infringement, it is insufficient to show copying from a protected work. The prima facie case of infringement also requires establishing that the copied work is substantially similar to the original.\textsuperscript{140} This improper appropriation requirement has been under pressure in recent decades.\textsuperscript{141} One line of decisions eroded the requirement by characterizing it as a rare exception for de minimis copying that can prevent liability only when the taking is so trivial as to be completely unrecognizable.\textsuperscript{142} For example, one recent decision found that the copying of a modified 0.23-second segment of horns from a protected song did not constitute improper appropriation only because “a reasonable juror could not conclude that an average audience would recognize the appropriation . . . .”\textsuperscript{143}

Other decisions have found improper appropriation on the ground that even the scant copied material at issue cleared the low hurdle of being recognizable as having been taken from the original.\textsuperscript{144} This is a

\textsuperscript{140} See Arnstein v. Porter, 154 F.2d 464, 468 (2d Cir. 1946) (identifying two elements necessary for a copyright infringement case: copying and improper appropriation); Nimmer & Nimmer, supra note 103, § 13.01[B] (explaining the elements a plaintiff must prove in an infringement action, including the substantially similar standard).

\textsuperscript{141} See Oren Bracha, Not De Minimis: (Improper) Appropriation in Copyright, AM. U. L. REV., at 1 (forthcoming 2018) (on file with authors) (discussing recent attacks on the requirement for improper appropriation as a prerequisite for copyright infringement).

\textsuperscript{142} See VMG Salsoul, LLC v. Ciccone, 824 F.3d 871, 887 (9th Cir. 2016) (“We hold that the ‘de minimis’ exception applies to actions alleging infringement of a copyright to sound recordings.”); Newton v. Diamond, 388 F.3d 1189, 1193 (9th Cir. 2003) (“[T]he de minimis maxim and the general test for substantial similarity . . . look[] to the response of the average audience, or ordinary observer, to determine whether a use is infringing.”); Tufenkian Imp./Exp. Ventures, Inc. v. Einstein Moomjy, Inc., 338 F.3d 127, 136 (2d Cir. 2003) (“[T]his is one of those relatively unusual cases in which the infringing work has copied the original . . . .”); Fisher v. Dees, 794 F.2d 432, 434 n.2 (9th Cir. 1986) (holding that de minimis standard is satisfied only if a taking “is so meager and fragmentary that the average audience would not recognize the appropriation”).

\textsuperscript{143} VMG Salsoul, 824 F.3d at 880. The court made its ruling with respect to both the copyright in the musical composition and the sound recording. Id. at 879–80.

\textsuperscript{144} See, e.g., Danjaq, LLC v. Universal City Studios, LLC, No. CV 14–02527 SJO (Ex), 2014 WL 7882071, at *5 (C.D. Cal. Oct. 2, 2014) (noting that “iconic phrases such as ‘Bond. James Bond’ are both
disconcerting development that conflicts with both the central role that
improper appropriation plays in modern copyright law as well as this
doctrine’s underlying policy.\footnote{145}

This enervation of the improper appropriation requirement is alarming.
Improper appropriation is one of copyright’s central mechanisms for striking
a balance between its beneficial incentives and social costs.\footnote{146} In particular,
the requirement limits copyright’s negative dynamic effect on future
creation by allowing beneficial secondary uses of works whose inclusion
within the owner’s exclusive rights would produce little marginal incentive
for creation.\footnote{147} Improper appropriation doctrine is thus a central mechanism
for recognizing that cultural innovation often happens through borrowing,
and even partial imitation of, existing works. More practically, this doctrine
helps balance the conflicting effects of copyright as both a promoter of
creation and a potential impediment to it.

A predictable objection is that the concern about the erosion of improper
appropriation is misguided because the fair use doctrine can adequately
carry out the function of shielding desirable secondary uses. Moreover, one
could argue that the preferred alternative is making fair use the primary
doctrinal lever for maintaining freedom for secondary uses. By reducing
improper appropriation to a rarely applied and relatively trivial exception,
this doctrinal design generally prevents any wasteful and confusing
duplication of coverage, and fair use’s open-endedness suggests that, in
principle, it is up to the desired protective task. Accordingly, some
commentators have criticized “the confusing overlap” between fair use and
the improper appropriation requirement,\footnote{148} and at least one has called for the
abolition of the latter on this ground.\footnote{149}

\footnote{145} See Bracha, supra note 141 (discussing recent attacks on satisfaction of the requirement for
improper appropriation as a prerequisite to finding copyright infringement).

\footnote{146} Cf. Benjamin Kaplan, An Unhurried View of Copyright 67 (Iris C. Geik et al. eds., 1967)
(criticizing the view “that infringement itself is decided without contamination by notions of
policy . . . ”).

\footnote{147} Bracha, supra note 141.

\footnote{148} Amy B. Cohen, Masking Copyright Decisionmaking: The Meaninglessness of Substantial
Similarity, 20 U.C. Davis L. Rev. 719, 745–46 (1987); see also Laura G. Lape, The Metaphysics of the
Law: Bringing Substantial Similarity Down to Earth, 98 Dickinson L. Rev. 181, 189 (1994) (referring
to some scholars noticing and criticizing the overlap between fair use and substantial similarity); Michael
Ferdinand Sitzer, Copyright Infringement Actions: The Proper Role for Audience Reactions in
Determining Substantial Similarity, 54 S. Cal. L. Rev. 385, 395 (1981) (criticizing “the duplication of
investigating market effects in the determination of substantial similarity” with the fourth fair use factor).

\footnote{149} Cohen, supra note 148, at 760–61 (calling for the abolition of improper appropriation and
considering of dissimilarities between works only as part of the fair use analysis in order to eliminate
As with similar objections to overlap between § 102(b) and fair use, this anti-redundancy objection falls flat because it fails to take seriously the way in which the two supposedly completely overlapping doctrines work in practice to advance relevant policies.\textsuperscript{150} Section II.A discussed one set of reasons why fair use alone is insufficient to secure appropriate breathing space for secondary creation.\textsuperscript{151} Being an open-ended standard applied by courts as an affirmative defense, fair use is often unpredictable and likely to require expensive and uncertain legal proceedings.\textsuperscript{152} The result is a chilling effect on the activities of parties who might prefer to steer clear of the gray area of uses of copyrighted works that a court might or might not deem fair.\textsuperscript{153}

Is improper appropriation any better than fair use in this respect? Arguably, the substantial similarity standard, which is applied case by case and often by juries, is at least as open-ended and unpredictable as fair use.\textsuperscript{154} Indeed, the question of whether there is substantial similarity is a wildcard in many copyright cases.\textsuperscript{155} Hence, in permitting doctrines of fair use and improper appropriation to overlap, we apparently do not derive the benefits of the rules-versus-standard overlay enabled through overlapping coverage of fair use and § 102(b). Without this, an anti-redundancy advocate might contend, there can be no justification for overlapping coverage of fair use

\textsuperscript{150} The analysis here assumes that the scope of the right extends well beyond verbatim reproduction as is true of modern copyright. It is this extended scope that necessitates a regulating mechanism that preserves breathing space for secondary uses. It has been suggested that the best alternative is to roll back copyright to its much narrower historical scope. See Rebecca Tushnet, \textit{Worth a Thousand Words: The Images of Copyright}, 125 HARV. L. REV. 683, 738 (2012) (proposing that we should abandon substantial similarity altogether). We bracket this option here for reasons of feasibility. This leaves fair use and improper appropriation as the main doctrinal tools for addressing the policy issues underlying copyright’s scope.

\textsuperscript{151} See supra text accompanying notes 106–16 (explaining the difficulty in applying, and predicting outcomes of, the fair use doctrine).

\textsuperscript{152} See supra notes 106–08 (discussing the unpredictability of outcomes in fair use cases).

\textsuperscript{153} See supra text accompanying notes 117–30 (noting the long and uncertain legal process in fair use cases, which can lead to overdeterrence).

\textsuperscript{154} See Peter Pan Fabrics, Inc. v. Martin Weiner Corp., 274 F.2d 487, 489 (2d Cir. 1960) (stating that the test for infringement of a copyright is necessarily “vague” and determinations must be made “ad hoc”). For scholarly criticism of the open ended nature of the standard of improper appropriation see Cohen, supra note 148, at 741–44 (“There is no objective framework for defining how much copying is too much.”); Lape, supra note 148, at 190–94 (discussing the difficulties associated with an amorphous standard); Mark A. Lemley, \textit{Our Bizarre System for Proving Copyright Infringement}, 57 J. COPYRIGHT SOC’Y U.S.A. 719, 719 (2010) (discussing the different approaches to the substantial similarity test); Jarrod M. Mohler, \textit{Toward a Better Understanding of Substantial Similarity in Copyright Infringement Cases}, 68 U. CIN. L. REV. 971, 988–89 (2000) (arguing that a single overarching test would further interests in consistency); see also Tushnet, supra note 150, at 716–17 (“The substantial similarity test is notoriously confusing and confused, perplexing students and courts alike.”).

\textsuperscript{155} See Joseph P. Fishman, \textit{Music as a Matter of Law}, 131 HARV. L. REV. 1861, 1905–07 (2018) (describing how substantial similarity is one of the most difficult questions in copyright law).
and the requirement of improper appropriation.

This objection also fails to give sufficient attention to the specific focus of the relevant doctrines and how they function in practice. While improper appropriation can be as unpredictable as fair use, at least in an important subset of the cases, a robust version of the doctrine is superior in generating more certainty ex ante and a simpler process ex post. Those are cases in which the conclusion that finding the uses infringing would impose substantial cost on future creation for little added marginal incentive can be reached relatively easily on the basis of limited information restricted to comparison of the two works and minimal context about them. The reason why there is a significant subset of such cases is the unitary, narrow focus of improper appropriation on such comparisons.

Recall that, in contrast, fair use is a multi-focal doctrine that concerns itself not only with how an alleged infringing work compares to copyrighted material, but also considers factors such as the purpose and market effect of alleged infringement. Fair use covers many different categories of circumstances under which a use should be deemed non-infringing for any of a variety of reasons. This capaciousness accounts for both the power and weakness of fair use. By contrast, improper appropriation is much more focused. The doctrine is designed to locate cases where, on the basis of comparison of the expressive content of two works and relatively minimal contextual information, it is possible to make a sound assessment that the cost of including the relevant secondary use within the exclusionary property right substantially outweighs its marginal incentive benefit.

Consider for example *Warner Bros. v. American Broadcasting Companies*, in which the character of Ralph Hinkley of the television show “The Greatest American Hero” was claimed to infringe the copyright in the character of Superman. Given the concept of the show as “what happens when you [the average person] become Superman,” it is unsurprising that the character of Hinkley had many traits reminiscent or evocative of Superman. In essence, however, Hinkley is an ordinary school

---


157 See supra note 53 and accompanying text (describing the four statutory factors courts use when considering the fair use doctrine).

158 See Joseph P. Liu, *Fair Use, Notice Failure, and the Limits of Copyright as Property*, 96 B.U. L. REV. 833, 837 (2016) (“As a contextual, multifactor standard, the fair use defense contains within it, inherently, a degree of uncertainty . . . .”).

159 See discussion supra Part II, Section B (discussing the scope of inappropriate appropriation and its role in modern copyright law).

160 720 F.2d 231 (2d Cir. 1983).

161 Id. at 235.

162 Id. at 236.
teacher who is forced reluctantly into superhero status. He uses his powers clumsily due to the loss of the instruction book, which is the antithesis of Superman. Recognizing this, the court of appeals in Warner Bros. upheld the trial court’s grant of summary judgment for the defendant due to the lack of substantial similarity, concluding, “[i]n the genre of superheroes, Hinkley follows Superman as, in the genre of detectives, Inspector Clouseau follows Sherlock Holmes.”

Could the same outcome have been reached on the basis of fair use? Probably. Indeed, the decision briefly discusses fair use doctrine but does not pursue it as a basis for non-liability. The point is that a fair use analysis in this case would have been more complex, costly, and open to unexpected twists and turns. The improper appropriation analysis required comparing the expressive content of the two works and the material difference in their meaning, notwithstanding the technical similarities. This relatively simple analysis served as a good proxy for the underlying policy question about the likely incentivizing effects of extending copyright protection to such uses, both on incentives to create original characters such as Superman and on subsequent creativity by others who wished to draw on established characters and their dominant cultural significance. Crucially, the analysis did not require a taxing process of discovery and fact finding.

In light of later developments in fair use case law, especially the rise in importance of the question of the transformative character of a secondary use, there is a good chance that today, the defendant in a case such as Warner Bros., would win on the basis of fair use. Still, the analysis would have to run through multiple tricky questions such as: (1) the weight to be imputed to the commercial nature of the use; (2) whether and the extent

---

163 Id.
164 Id.
165 Id. at 243.
166 Id. at 242–43.
167 Id. at 243.
168 See Peter F. Gaito Architecture, LLC v. Simone Dev. Corp., 602 F.3d 57, 64 (2d Cir. 2010) (“When a court is called upon to consider whether the works are substantially similar, no discovery or fact-finding is typically necessary.”); Nimmer & Nimmer, supra note 103, § 12.10[B][3] (“[T]he works themselves supersede and control any contrary allegations, conclusions, or descriptions of the works contained in the pleadings.”).
169 See Campbell v. Acuff-Rose Music Inc., 510 U.S. 569, 587 (1994) (“[W]hether ‘a substantial portion of the infringing work was copied verbatim’ from the copyrighted work is a relevant question for it may reveal a dearth of transformative character. . . .”’) (quoting Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S., 565 (1985)); Cariou v. Prince, 714 F.3d 694, 706–07 (2d Cir. 2013) (discussing transformative use and how it is critical to consider what a reasonable observer would say about a particular work, not just what an artist might say); Bill Graham Archives v. Dorling Kindersley, Ltd., 448 F.3d 605, 612 (2d Cir. 2006) (finding that the defendant was using the plaintiff’s images for a transformative use).
to which the use could be considered transformative or even parodic;\textsuperscript{171} (3) the amount and substantiality of the material taken in relation to the purpose of the use;\textsuperscript{172} and (4) the likely market effect on profits from the original in light of lack of an apparent market failure in negotiating a license on the one hand\textsuperscript{173} and the possible transformative character of the use on the other.\textsuperscript{174} Each of these questions could precipitate not only legal argument, but also demands for discovery and fact finding,\textsuperscript{175} with the assenter of the affirmative defense of fair use bearing the burden of establishing a basis for favorable conclusions on factors such as market effect.\textsuperscript{176} Hence, even where fair use analysis would lead to the same no-liability conclusion as a robust improper appropriation requirement, improper appropriation doctrine can provide a much more effective, less costly way of securing social interests in activities that only partially draw on expression in preexisting works.

The narrower focus of improper appropriation does not mean that there will be no hard cases. It does mean, however, that there will be a core set of cases where the application of a robust version of this doctrine should be a relatively simple and certain way of determining non-liability, whereas fair use will at best offer only a tortuous and, for many parties, overly uncertain path to the same end. As a result, there will be an important set of cases where a robust improper appropriation requirement can more cheaply and

\textsuperscript{171} Campbell, 510 U.S. at 579–81. Compare Bill Graham Archives, 448 F.3d at 608–12 (finding use of concert posters in a biographical book to be transformative), and Suntrust Bank v. Houghton Mifflin Co., 268 F.3d 1257, 1269–71 (11th Cir. 2001) (finding the extensive use of materials from original novel to be transformative and parodic), with Salinger v. Colting, 641 F. Supp. 2d 250, 256–63 (S.D.N.Y. 2009) (finding the use of materials from original novel neither parodic nor highly transformative), vacated, 606 F.3d 68 (2d Cir. 2010), and TCA Television Corp. v. McCollum, 839 F.3d 168, 180–82 (2d Cir. 2016) (finding the use of a comedy routine in a play not transformative).

\textsuperscript{172} 17 U.S.C. § 107(3) (2012); Campbell, 510 U.S. at 586–89.

\textsuperscript{173} 17 U.S.C. § 107(4) (2012); see Am. Geophysical Union, 60 F.3d at 930 (noting that only “traditional, reasonable, or likely to be developed markets” are to be considered under the fourth fair use factor). In Warner Bros., there was no market failure in a conventional sense because prior to the unauthorized use the parties engaged in licensing negotiations. See Warner Bros. v. Am. Broad. Cos., 720 F.2d 231, 236 (2d Cir. 1983) (noting ABC’s failure in obtaining a production license for a television series about “Superboy,” a series based on the early adventures of Superman).

\textsuperscript{174} See Campbell, 510 U.S. at 591 (“[W]hen . . . the second use is transformative . . . market harm may not be so readily inferred.”); see also Bill Graham Archives, 448 F.3d at 614 (explaining that licensing markets for transformative uses are not considered under the fourth fair use factor even if already established).

\textsuperscript{175} Demonstrative of this point is the fact that in Campbell, which is seen as a strong pro-fair-use decision, the Supreme Court refrained from deciding the issue on the merits and remanded the case to the trial court with instructions for further proceedings, presumptively to include further fact finding, on the issues of the substantiality of the copyrighted material taken and the effect on the relevant secondary market. 510 U.S. at 593–94. An appellate court may decide fair use as a matter of law only when the trial court “has found facts sufficient to evaluate each of the statutory factors.” Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 560 (1985).

\textsuperscript{176} See Campbell, 510 U.S. at 590 (describing the evidentiary presumption applied by the Court of Appeals).
predictably shield beneficial secondary uses, and thus more effectively serve the aims of underlying copyright policy.

The substantive superiority of improper appropriation for an important set of cases is boosted by procedural aspects of litigation.\textsuperscript{177} As indicated by concerns with litigation costs highlighted in Section I.A, whether a legal rule can effectively shield desirable secondary uses of copyrighted works does not just depend on the substantive result generated by the rule post-litigation or even the predictability of that result. Restricting vindication of a privilege to use to a procedural route requiring long and expensive litigation increases the risk of undesirable chilling effects on secondary uses.\textsuperscript{178} Conversely, a simpler and more accessible path for vindicating the privilege is likely to both boost the willingness of users to take the risk and even decrease threats from copyright holders issued in the shadow of litigation.\textsuperscript{179}

Summary judgment can act as a key mechanism here by enabling a user to obtain vindication of its privilege to use without the need for a full trial. Courts have shown some willingness to grant summary judgment to defendants on the basis of findings of fair use in the appropriate cases.\textsuperscript{180} Nonetheless, by providing a route to decisions that rely on a more focused inquiry than fair use, a robust improper appropriation doctrine makes summary judgment in favor of a secondary user significantly easier for a court to grant in an important subset of cases.\textsuperscript{181} This feature of the improper appropriation doctrine is bolstered by the fact that, unlike fair use, improper appropriation is an element of the prima facie case for which the burden of

\textsuperscript{177} For a more detailed discussion of this point, see Bracha, supra note 141.
\textsuperscript{178} See supra text accompanying note 116 (discussing the problem of overdeterrence in copyright).
\textsuperscript{180} See Wright v. Warner Books, Inc., 953 F.2d 731, 735 (2d Cir. 1991) (noting that courts have resolved fair use determinations at the summary judgment stage); Cable/Home Commc’n Corp. v. Network Prods. Inc., 902 F.2d 829, 845 (11th Cir. 1990) (concluding that plaintiff’s copyright infringement claim warranted summary judgment because defendant’s affirmative defense of fair use was unsupported); Fisher v. Dees, 794 F.2d 432, 440 (9th Cir. 1986) (holding that the parody song in question is deserving of fair-use protection as a matter of law); see also NIMMER & NIMMER, supra note 103, § 12.10[B][4]. For a discussion of the rules pertaining to summary judgment and fair use, see Ned Snow, Fair Use as a Matter of Law, 89 DENV. L. REV. 1, 2 (2011).
\textsuperscript{181} See Durham Indus., Inc. v. Tomy Corp., 630 F.2d 905, 918 (2d Cir. 1980) (“[W]here both the plaintiff’s and defendant’s works are before the court, ‘the court may compare the two works and render a judgment for the defendant on the ground that as a matter of law a trier of fact would not be permitted to find substantial similarity.’” (citation omitted)); Malibu Media, LLC v. Doe, 82 F. Supp. 3d 650, 653 n.1 (E.D. Pa. 2015) (“Upon review of the works themselves, if the court concludes that no trier of fact could rationally determine the two to be substantially similar, it can render a defense judgment as a matter of law.”); Wavelength Film Co. v. Columbia Pictures Indus., Inc., 631 F. Supp. 305, 307–08 (N.D. Ill. 1986) (concluding that no reasonable jury could find the two disputed films to be substantially similar); Anderson v. Paramount Pictures Corp., 617 F. Supp. 1, 3 (C.D. Cal. 1985) (holding that there was no genuine issue that defendants infringed plaintiff’s copyright); see also NIMMER & NIMMER, supra note 103, §12.10[B][3].
proof is on the plaintiff. For example, the Warner Bros. case discussed above was decided by summary judgment on the basis of a comparison of the original and allegedly infringing works with little resort to other factual concerns. Consider also the case of Walker v. Time Life Films, Inc., where the trial court granted summary judgment rejecting an infringement claim by an author against the producers and screenwriter of a motion picture that was allegedly based on the author’s book. The court reached its decision after analyzing the two works on the basis of a thin version of stipulated facts agreed to by the parties for purposes of the motion, as well as expert opinions analyzing the works.

This divergence between the general scope and evidentiary demands of inquiries into improper appropriation and fair use is even more important in relation to prospects for terminating a meritless case even earlier through a motion to dismiss. Unlike fair use, improper appropriation is part of the prima facie case of copyright infringement that a plaintiff must establish. This is important in light of two other factors: (1) the current pleading burden requiring a complaint in a civil case to plausibly suggest an entitlement to relief, and (2) the willingness of some courts to grant motions to dismiss in copyright cases due to failure to establish improper appropriation. Together, these factors mean that courts can dismiss weak infringement claims at a very early stage—even prior to discovery—on the basis of the pleadings and a comparison of the expressive content of the

---

182 Nimmer & Nimmer, supra note 103, § 13.01[B]. Courts demand more from a party seeking summary judgment on the basis of a claim for which that party bears the burden of proof in trial. See 10A Charles Alan Wright, Arthur R. Miller & Mary Kay Kane, Federal Practice and Procedure § 2727.1 (4th ed. 2010) (“Rule 56 first imposes a burden of production on the moving party to make a prima facie showing that it is entitled to summary judgment.”).


184 615 F. Supp. 430 (S.D.N.Y. 1985), aff’d, 784 F.2d 44 (2d Cir. 1986).

185 Id. at 432–33.

186 See id. at 433 n.1 (noting that the facts were provided to the court through statements submitted by the parties).

187 Id. at 438; Walker, 784 F.2d at 47–48.


189 See supra note 182 and accompanying text (explaining that the burden of production is first imposed on the moving party to make a prima facie showing that it is entitled to summary judgment).

190 Bell Atlantic Corp. v. Twombly, 550 U.S. 544, 557 (2007) (holding that a complaint requires allegations “plausibly suggesting” a valid claim).

191 See Peter F. Gaito Architecture, LLC v. Simone Dev. Corp., 602 F.3d 57, 64 (2d Cir. 2010) (noting that the court may determine noninfringement on a motion to dismiss when the similarity concerns only noncopyrightable elements of the plaintiff’s work or when no reasonable trier of fact could find the works substantially similar); Nelson v. PRN Prods., Inc., 873 F.2d 1141, 1143–44 (8th Cir. 1989) (holding that the district court properly granted the defendant’s motion to dismiss on the ground that it failed to state a claim for infringing use); see also Nimmer & Nimmer, supra note 103, § 12.10[B][3] (discussing how common stock with a particular expressive genre do not receive copyright protection).
works in the record.¹⁹² This avenue for prompt dismissal, which hardly seems available under fair use,¹⁹³ bolsters the improper appropriation doctrine’s capacity—and differential capacity compared to fair use—to discourage, or otherwise lead to early disposal of, especially dubious claims of copyright infringement.¹⁹⁴

The provision of a relatively fast and cheap way of dismissing meritless infringement suits is of particular importance when one widens the normative lens beyond economic efficiency. As a field regulating and shaping expressive activities, copyright is of concern to several normative theories focused on self-determination, the democratic public-sphere, or human flourishing.¹⁹⁵ For our purposes, the crucial, common feature of these

¹⁹² See, e.g., Rentmeester v. Nike, Inc., 883 F.3d 1111, 1123, 1125 (9th Cir. 2018) (affirming a grant of a motion to dismiss for lack of improper appropriation and concluding that “[n]othing disclosed during discovery could alter the fact that the allegedly infringing works are as a matter of law not substantially similar” to the plaintiff’s work); Gaito Architecture, 602 F.3d at 64 (“[N]o discovery or fact-finding is typically necessary, because ‘what is required is only a visual comparison of the works.’”) (citing Folio Impressions, Inc. v. Byer Cal., 937 F.2d 759, 766 (2d Cir. 1991)).

¹⁹³ See, e.g., Nelson, 873 F.2d at 1144 (affirming the trial court’s decision to grant motion to dismiss without allowing further discovery because additional information was relevant for refuting a fair use defense but not for substantial similarity).

¹⁹⁴ An additional possible source of support for the need of a readily available mechanism for dismissal of meritless copyright suits is the phenomenon colloquially referred to as “copyright trolls.” These are entities in the business of enforcing copyrights in which they acquire an interest. Allegedly, this activity often exploits various aspects of copyright law in order to extract settlement payments even when the legal claim is dubious, which had been traditionally under-enforced or is of limited economic impact. In recent years, copyright scholarship has started taking note of this phenomenon. See, e.g., Shyamkrishna Balganesh, The Uneasy Case Against Copyright Trolls, 86 S. CAL. L. REV. 723, 728–29 (2013) (arguing that the delegation of copyright enforcement is troublesome because it disrupts the balance of copyright claims that otherwise would have been tolerated); James DeBriyn, Shedding Light on Copyright Trolls: An Analysis of Mass Copyright Litigation in the Age of Statutory Damages, 19 UCLA ENT’L. L. REV. 79, 86–90 (2012) (discussing the different business models of copyright trolling throughout the last century); Brad A. Greenberg, Copyright Trolls and Presumptively Fair Uses, 85 COLO. L. REV. 53, 53 (2014) (explaining that the fair use doctrine is copyright law’s internal limitation on the enforcement-only business model of “copyright trolling”); Matthew Sag, Copyright Trolling, An Empirical Study, 100 IOWA L. REV. 1105, 1105 (2015) (analyzing the correlation between copyright trolling and the growth of multi-defendant John Doe litigation over the past decade); Samuelson, supra note 1, at 759 (discussing how copyright trolls likely generate litigation disproportionate to the harm caused). Courts too have started referring to copyright trolls in their decisions. See, e.g., Brownmark Films, LLC v. Comedy Partners, 682 F.3d 687, 691–92 (7th Cir. 2012) (stating that the plaintiff had “the appearance of a ‘copyright troll’”); Third Degree Films v. Does 1–47, 286 F.R.D. 188, 189 (D. Mass. 2012) (“[T]he Court has grueously troubled by ‘copyright trolling . . . .’”). There is, however, considerable disagreement on the extent and character of the phenomenon as well as its evaluation. See Shyamkrishna Balganesh & Jonah B. Gelbach, Debunking the Myth of the Copyright Troll Apocalypse, 101 IOWA L. REV. 43, 44 (2016) (analyzing empirical observations about the phenomenon of copyright trolling). Note that the case for the importance of an efficient mechanism for the early and cheap dismissal of meritless copyright lawsuits does not hinge on the anti-copyright-trolls argument, but is only bolstered by it.

theories is that, in evaluating copyright’s institutional design, they all go beyond strict comparison of positive and negative individual welfare effects. Each of these normative frameworks, for its own reasons, places a premium on a robust freedom to engage in secondary uses of copyrighted works. Each also takes an interest not just in a bottom line for overall social welfare, but also in questions of which members of society get to engage in which kinds of expressive activities. Because these theories are interested in broad and diverse access to secondary uses for creative purposes (“secondary creation”), they provide strong justifications for improper appropriation’s procedural advantages. As discussed above in relation to fair use, the chilling effect of dubious claims is, in practical effect, borne unequally by different secondary creators depending on their sophistication and financial backing. Hence, by providing greater opportunities for quick and cheap dismissal of such claims, improper appropriation doctrine is likely to reduce differential chilling effects and thus to increase both the number and diversity of actors who engage in secondary creation.

An additional counter-argument to the anti-redundancy objection is focused not on the superiority of improper appropriation over fair use in shielding certain secondary uses, but on the beneficial character of the overlap between the two. A doctrinal structure that includes both doctrines more robustly enables secondary uses because it creates a double barrier for liability. To establish copyright infringement, a plaintiff must both establish improper appropriation and survive any assertion of a fair use defense. A defendant can prevail by either a plaintiff’s failure to establish improper appropriation or a successful fair use defense. One may be tempted to say that redundancy functions here as an error-check: a court that gets wrong the application of one of the doctrines can still get to the right outcome under the other. This is true, but only if the correct result is no liability. When the correct result is infringement—a result that requires sub-results under both doctrines to point in this direction—the double barrier can be expected to increase the risk of error rather than reduce it. Hence, part of the overall policymaking question becomes whether we should be more worried about “Type I” errors that improperly subject secondary uses to copyright liability—false positives—or “Type II” errors that improperly excuse such

196 See id. at 251–58 (describing four normative theories of copyright and analyzing how each engages with expression).
197 Id. at 258–87.
198 See supra notes 116–130 and accompanying text (discussing the chilling effect of uncertainty in application of the fair use defense).
199 See Liu, supra note 116, at 434 (“[A] chilling effect may result from the often significant economic disparity between copyright owners and those who seek to build upon pre-existing works.”); Molly Shaffer Van Houweling, Distributive Values in Copyright, 83 Tex. L. Rev. 1535, 1566 (2005) (“‘Get a license or do not sample’ is no answer to a would-be creator who does not have the money to participate in the license marketplace.”).
secondary uses—false negatives. At least at the level of a first-order analysis, the layering of a meaningful improper appropriation doctrine on top of a background defense of fair use favors avoiding Type I errors at the risk of increasing Type II errors.

Why might we want such an asymmetric structure of redundancy? One possible justification is that such a structure helps ensure that copyright law properly respects interests in freedom of speech. It is widely recognized that copyright and free speech have a natural tendency to conflict. By definition, copyright imposes limitations on speech that are backed by legal sanctions and remedies. Relative to the value of such limitations in promoting creative speech, the resulting discouragement of certain forms of speech and potential depopulation of the “marketplace of ideas” may be viewed as particularly, although not exclusively, problematic in cases where there exist significant substantial similarity questions—in other words, cases where the user’s expression materially differs from mere copying or repetition of another’s speech. Even more to the point with respect to redundancy, freedom of speech concerns are likely to be significant in cases implicating the doctrinal layering of an improper appropriation requirement and a fair use defense because these cases typically involve substantial transformation of original expression.

Courts and commentators often dismiss concerns about tension between

---

200 See, e.g., C. Edwin Baker, First Amendment Limits on Copyright, 55 VAND. L. REV. 891, 892 (2002) (“A legislative grant of this private power to stop speech on the basis of its content is in overt tension with the constitutional guarantees of speech and press freedom.”); Robert C. Denicola, Copyright and Free Speech: Constitutional Limitations on the Protection of Expression, 67 CAL. L. REV. 283, 284 (1979) (“Even a cursory examination of the origins of copyright law reveals the potential conflict between property rights in intellectual creations and freedom of expression.”); Neil Weinstock Netanel, Locating Copyright Within the First Amendment Skein, 54 STAN. L. REV. 1, 7 (2001) (“Copyright’s speech encumbrance cuts a wide swath . . . .”); Melville B. Nimmer, Does Copyright Abridge the First Amendment Guarantees of Free Speech and Press?, 17 UCLA L. REV. 1180, 1181 (1970) (posing the question of whether the Copyright Act “abridges the ‘freedom of speech’ and ‘of the press’ in that it punishes expressions by speech and press when such expressions consist of the unauthorized use of material protected by copyright”).

201 See Fogerty v. Fantasy, Inc., 510 U.S. 517, 527 (1994) (“Because copyright law ultimately serves the purpose of enriching the general public through access to creative works, it is peculiarly important that the boundaries of copyright law be demarcated as clearly as possible.”).

202 Cf. Joseph Blocher, Institutions in the Marketplace of Ideas, 57 DUKE L.J. 821, 829 (2008) (“Holmes’s invocation of the ‘marketplace of ideas’ metaphor, though it came in a dissent, has had as major an impact as any Supreme Court decision on popular and academic thinking about the First Amendment.”).


204 Cf. id. at 550 (discussing the view that transformative uses “that add[] value to the original br[ing] something new and creative into the world, so they might be justified [relative to copyright law’s purposes] even if the copyright owner objected”).

205 Netanel, supra note 111, at 736 (“[F]air use doctrine today is overwhelmingly dominated by the Leval-Campbell transformative use doctrine.”).
copyright and freedom of speech by observing that copyright incorporates internal mechanisms, such as the fair use doctrine and the rule that ideas are not copyrightable, which balance speech interests with those served by copyright. But doctrinal balancing of speech interests in principle does not necessarily translate into effective protection of those interests “in action.”

An overlap between the fair use doctrine and the requirement of improper appropriation, with the greater practical opportunities for expeditious adjudication that this overlap affords, is one way of helping ensure that copyright indeed takes free speech rights seriously. Under this view, a somewhat lopsided doctrinal design in favor of preventing errors in findings of liability is a feature, not a bug.

Particularly in cases where a user of copyrighted material has added new expressive content, double-layered protection of the user’s freedom of speech can make perfect sense. Limiting false positives in the form of findings of liability can be viewed as more important than limiting false negatives in the form of failures to find liability when a fundamental interest in free speech is aligned with the former goal. This justification of the redundant structure of fair use and improper appropriation fits a long tradition of procedural safeguards for speech and of limiting the reach of

---

206 See Eldred v. Ashcroft, 537 U.S. 186, 221 (2003) (stating that “copyright’s built-in free speech safeguards are generally adequate to address” First Amendment concerns); Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 555–60 (1985) (declining to expand “the doctrine of fair use to create what amounts to a public figure exception to copyright”); New Era Publ’ns Int’l, ApS v. Henry Holt & Co., 873 F.2d 576, 584 (2d Cir. 1989) (asserting that the Second Circuit’s prior “observation that the fair use doctrine encompasses all claims of first amendment in the copyright field never has been repudiated” (internal citation omitted)); Sid & Marty Krofft Television Prods., Inc. v. McDonald’s Corp., 562 F.2d 1157, 1170 (9th Cir. 1977) (arguing that the idea/expression dichotomy accommodates competing copyright and First Amendment interests). This argument is usually traced to Nimmer, supra note 200, at 1181. The built-in speech balancing mechanisms commonly highlighted are the idea/expression dichotomy, the fair use doctrine, and copyright’s limited duration.

207 See David S. Olson, First Amendment Interests and Copyright Accommodations, 50 B.C. L. Rev. 1393, 1413 (2009) (“[D]ue to the significant changes to the traditional contours of copyright . . . the idea/expression dichotomy and fair use doctrines cannot come close to adequately protecting the public’s interests in speech . . . .”); Roscoe Pound, Law in Books and Law in Action, 44 AM. L. Rev. 12, 36 (1910) (“It is the work of lawyers to make the law in action conform to the law in the books . . . . by making the law in the books such that the law in action can conform to it, and providing a speedy, cheap and efficient legal mode of applying it.”); see also Netanel, supra note 200, at 26 (“[I]t is highly—and increasingly—doubtful that copyright limitations adequately protect First Amendment values.”); Tushnet, supra note 203, at 554 (“[F]air use, with its balancing apparatus, is ill suited to protecting activities that are at the core of ordinary uses of copyrighted works; it is supposed to deal with unusual or marginal activities.”).

208 Cf. RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 186 (1977) (“If we cannot insist that the Government reach the right answers about the rights of its citizens, we can insist at least that it try.”).

private law doctrines to accommodate free speech interests.\textsuperscript{210} Bracketing the question of whether this doctrinal structure can immunize copyright from First Amendment challenges in any particular case, the one-sided redundancy seems an important mechanism for erring on the side of caution when deciding whether to burden speech with copyright liability.

\textbf{C. Copyrightable Subject Matter and Originality}

Section 102(a) of the Copyright Act provides that copyright “subsists, in accordance with this title, in original works of authorship.”\textsuperscript{211} On its face, this section sets forth two requirements for copyrightable subject matter: (1) it must be a “work[] of authorship”; and (2) it must be original.\textsuperscript{212} As indicated in Section II.A,\textsuperscript{213} § 102(b) adds a further set of limitations by explicitly excluding from coverage “any idea, procedure, process, system, method of operation, concept, principle, or discovery.”\textsuperscript{214} The Act does not expressly define “works of authorship,” but § 102(a) does provide a non-exhaustive list of eight categories.\textsuperscript{215} Copyright’s originality requirement, for its part, is commonly understood to require independent creation by the author\textsuperscript{216} and “more than a de minimis quantum of creativity.”\textsuperscript{217}

The net result of this doctrinal structure is partial overlap between three operative parts: (1) originality; (2) the general copyrightable subject-matter requirement; and (3) express subject-matter exclusions. Tricky questions can result when a decision-maker has to assess the copyrightability of subject matter that is at best a poor fit for the statute’s illustrative categories. The statute’s apparent use of “original” as a limiting modifier for the separate term “works of authorship” suggests that Congress envisioned the originality and subject-matter requirements as imposing distinguishable limitations.\textsuperscript{218}

\textsuperscript{210} See New York Times Co. v. Sullivan, 376 U.S. 254, 292 (1964) (“We hold today that the Constitution delimits a State’s power to award damages for libel in actions brought by public officials against critics of their official conduct.”).

\textsuperscript{211} 17 U.S.C. § 102(a) (2012).

\textsuperscript{212} Id.

\textsuperscript{213} See supra Section II.A (exploring limitations imposed by § 102(b)).

\textsuperscript{214} 17 U.S.C. § 102(b) (2012).

\textsuperscript{215} 17 U.S.C. § 102(a) (2012); see also Nimmer & Nimmer, supra note 103, § 2.03[A] (discussing the meaning of the phrase “works of authorship” in 17 U.S.C. § 102(a) by pointing to the statutory language and a related congressional report).

\textsuperscript{216} See Nimmer & Nimmer, supra note 103, § 2.01[A][1] (stating that an original work is one that is the “product of the independent efforts of its author”).


\textsuperscript{218} 17 U.S.C. § 102(a) (2012).
Nonetheless, courts sometimes reduce the analysis to an inquiry into originality and marginalize or ignore altogether the independent question of whether relevant subject matter is a work of authorship within the statutory ambit. An example of this reductive approach is the 2005 decision in *Open Source Yoga Unity v. Choudhury*.\(^{219}\) In this case, the district court assessed the copyrightability of a sequence of yoga poses and, in so doing, appeared to collapse both § 102(a)’s “work of authorship” requirement and § 102(b)’s exclusions into the originality requirement.\(^{220}\) The court responded to the plaintiff’s argument that the yoga poses were “merely uncopyrightable functional physical movements” by holding that “if the trier of fact determines that a sufficient number of the individual yoga asanas are arranged in a sufficiently creative manner, copyright protection for the yoga sequence would be available.”\(^{221}\) Thus, the originality analysis completely swallowed any independent analysis of whether the yoga sequences qualified as a “work of authorship” and whether they were excluded as a functional “procedure.”\(^{222}\) In this way, the district court perhaps unintentionally adopted a position analogous to that of patent law commentators who contend that the patentability of a claimed invention should turn essentially entirely on requirements of utility, novelty, non-obviousness, and adequacy of the patent’s written description, with any inquiry into the patent-eligibility of the claimed invention’s subject matter being no more than trivial.\(^{223}\)

The Copyright Office has taken a different view. First, the Office has stated that originality does not establish subject-matter eligibility; according to the Office, “a creative selection, coordination, or arrangement must also result in one or more congressionally recognized categories of authorship.”\(^{224}\) Second, the Office has taken the more controversial\(^{225}\) step

\(^{219}\) 74 U.S.P.Q.2d 1434 (N.D. Cal. 2005).

\(^{220}\) Id. at 1436–38 (providing the court’s analysis of “Copyright Validity” of a sequence of yoga asanas).

\(^{221}\) Id. at 1437–38

\(^{222}\) See id. at 1437 (stating that issue of originality is the “backdrop” against which the claims “must be evaluated”).


\(^{225}\) There is a plausible argument that the § 102(a) categories are illustrative and leave the door open for judicial recognition of other works of authorship. See Nimmer & Nimmer, *supra* note 103, § 2.03[A] (noting that it is “clear that ‘works of authorship’ are not necessarily limited to the eight broad categories of works listed under Section 102(a)”) (quoting 17 U.S.C. § 102(a) (2012)); R. Anthony Reese, *Copyrightable Subject Matter in the “Next Great Copyright Act.”* 29 BERKELEY TECH. L.J. 1489, 1499 (2014) (observing that “the statutory term ‘works of authorship’ can include works that do not fall within any of the categories enumerated in section 102(a),” but recommending as a reform that “the statute should expressly enumerate all of the categories of works to which Congress chooses to grantcopyright protection and should not extend protection to an open-ended and undefined category of works.”);
of contending that neither the courts nor the Office may recognize “new categories of authorship” beyond the forms explicitly listed in the Act. On this basis, the Copyright Office has rejected “the copyrightability of the selection and arrangement of preexisting exercises, such as yoga poses,” except to the extent that such a compilation fits at least one of the statutorily listed categories. For example, a sequence of asanas would qualify as a work of “choreographic authorship” if “it contains sufficient attributes of a work of choreography” rising beyond the level of “social dance steps and simple routines” that legislative history indicates Congress meant to exclude from copyrightability. But unless the sequence fits this or another statutory category, it cannot be protected.

A recent decision by the Ninth Circuit addressed the same sequences of yoga poses as Open Source but took a third route. In Bikram’s Yoga College of India, L.P. v. Evolution Yoga, L.L.C., the court emphasized that passing the originality bar is a distinct matter from surviving the § 102(b) exclusion of functional subject matter. The court then determined that the asana was functional and therefore uncopyrightable because of health and psychological benefits touted by its creator. This outcome, the court concluded, made it unnecessary to examine the separate § 102(a) challenge.

How should these three doctrines work together? Are some of them superfluous in light of the others? Should courts, like that in Open Source, treat originality as effectively the test for whether there is a copyrightable work of authorship? Should the § 102(b) exclusion of nonfunctional subject matter be considered avoided whenever some aspect of a work satisfies the § 102(a) requirement of originality?

Pamela Samuelson, Evolving Conceptions of Copyright Subject Matter, 78 U. Pitt. L. Rev. 17, 17 (2016) (observing that un-enumerated categories of “other types of intellectual creations may be eligible for protection as long as they satisfy copyright’s originality and fixation requirements”).

Statement of Policy, supra note 224, at 37,607.

Id. (quoting H.R. Rep. No. 94-1476, at 54 (1976)).

Id. at 37, 607.

Bikram’s Yoga Coll. of India, L.P. v. Evolution Yoga, L.L.C., 803 F.3d 1032, 1036 (9th Cir. 2015).

See id. at 1038 (noting a court’s prior explanation that certain “‘processes , even if original, cannot be protected by copyright’” (quoting Palmer v. Braun, 287 F.3d 1325, 1334 (11th Cir. 2002))).

Id. at 1040 (“‘[B]eauty is not a basis for copyright protection.’”); see also Patry, supra note 52, § 4.22 (arguing that yoga sequences cannot be protected by copyright when functional).

Bikram’s Yoga, 803 F.3d at 1043 (“In this case, we need not decide whether to adopt the Copyright Office’s definition of ‘choreographic work’ or fashion another on our own . . . .”).

This way of avoiding the § 102(b) exclusions might be considered comparable to the European prohibition of patenting business methods being avoided when an applicant establishes that the claimed business method has “technical effect” and thus that the patent is not seeking to cover a bare business method per se. See Rajnish Kumar Rai & Srinath Jagannathan, Do Business Method Patents Encourage Innovation?, 2012 B.C. Intell. Prop. & Tech. F. 1, 6 (“[E]ven though the European patent system is
subject matter qualification under either subsection of § 102 is doubtful, the relevant work may also fail the originality bar. We believe, however, that even if there is some overlap between copyright’s subject matter requirements and the originality threshold, it makes substantial policy sense to have meaningful subject matter requirements that are separate and independent of copyright’s originality threshold. Both § 102(a) and § 102(b) have distinct functions that are far from superfluous.

Consider first the particular relationship between originality and the § 102(b) exclusion of functional subject matter. Copyright law includes various features, such as a (roughly) century-long copyright term and lax substantive and procedural thresholds for obtaining rights,\(^{235}\) that, from the standpoint of promoting social welfare, are plausibly a much better fit for some forms of subject matter than others. In the analogous case of patent law, exclusion from patentability of laws of nature and abstract ideas—however useful, novel, or nonobvious—can make sense because of concern that the patenting of such scientific and technological building blocks can create excessive bottlenecks for follow-on research and development.\(^{236}\) Similarly, some functional subject matter is a poor fit for copyright, although perhaps sometimes a decent fit for patent. A new machine or a system for calculating temperatures are not the kind of subject matter for which copyright’s doctrinal features are designed. Whether they are protected and what kind of protection they get should be a matter of patent law, which is designed for this kind of subject matter. Thus, the § 102(b) exclusion of functional subject matter serves as a first rough filter that improves the tailoring of copyright’s coverage to its social ends and does so in a reasonably administrable, or even administration-aiding, way. In other words, § 102(b) serves as a channeling mechanism. It points relevant subject matter to the legal regime appropriate for it on the basis of a rough categorical classification.\(^{237}\)

considered to be far more stringent . . . patent applications that describe the ‘technical effect’ of these methods in addition to claiming abstract business methods can obtain protection under the European Patent system.”).

\(^{235}\) See Kevin Emerson Collins, Patent Law’s Authorship Screen, 84 U. CHI. L. REV. 1603, 1615–16 (2017) (“When functional innovation infiltrates the copyright regime, copyright can override patent’s limitations on both the ability to obtain rights and the length of protection.”).

\(^{236}\) See John M. Golden, Patentable Subject Matter and Institutional Choice, 89 TEX. L. REV. 1041, 1069–74 (2011) (arguing that patents for “general scientific or mathematical ideas” will increase the costs and decrease the incentives for innovation).

\(^{237}\) See Baker v. Selden, 101 U.S. 99, 102 (1879) (“To give to the author of the book an exclusive property in the art described therein . . . is the province of letters-patent, not of copyright.”); Mark P. McKenna & Christopher Jon Sprigman, What’s In, and What’s Out: How IP’s Boundary Rules Shape Innovation, 30 HARV. J.L. & TECH. 491, 531 (2017) (arguing that § 102(b) establishes boundaries between copyright, patent, and the public domain); Pamela Samuelson, Strategies for Discerning the Boundaries of Copyright and Patent Protections, 92 NOTRE DAME L. REV. 1493, 1496–1500 (2017) (examining the constitutional and statutory reasons for drawing “sharp distinctions” between the domains of copyright and patent); Buccafusco & Lemley supra note 38, at 1304 (observing that “[f]unctionality
The relationship between the originality requirement and § 102(b) depends on the purposes served by the doctrines, which in turn should inform how the doctrines are understood and applied. Of the two distinct elements of the originality requirement—“independent creation” and a quantum of creativity—"independent creation"—the first one, demanding that the protected subject matter originate with the person claiming protection, is of little relevance here. Opinions vary on the proper purpose of the second element of the originality requirement, the requirement of a modicum of creativity of the protected subject matter.

The dominant justification for the very low creativity bar in existing law is avoidance of cumbersome evidentiary inquiries in establishing copying. According to this view, because originality imposes a threshold requirement that copyrightable material be distinguishable from preexisting material, courts are commonly spared the complicated task of figuring out whether the similarity of an allegedly infringing work is attributable to copying from the protected work or to both works having drawn on common sources.

For others who advocate a more demanding originality threshold, one that would require substantial revision of existing doctrine, the fundamental justifications for the requirement tend to differ. These alternative justifications include the following: (1) a better cost-benefit balance achieved by restricting protection only to individual cases where the creative contribution and hence the social benefit of the work are likely to be significant; (2) avoiding wasteful rent dissipation due to the creation of duplicative works seeking to capture valuable demand; and (3) ensuring
the conditions for human flourishing in the cultural sphere by focusing copyright protection on works with a creative aspect that makes them particularly likely to enrich culture.243

Under any of these justifications for the originality requirement, either the existing one or a revised version of it, the stated operational logic differs from that of § 102(b). In all of these variants, originality serves a different purpose from that of § 102(b) and specifies different desiderata, whether it is novelty, innovation, or evidentiary distinguishability. As a result, one might expect to see little redundancy or even overlap between the two doctrines.

There is, however, another understanding of the existing low creativity bar that construes it as having more overlap with § 102(b). Under this understanding, a purpose of this minimal creativity requirement is to ensure that copyright is only extended to the kind of subject matter for which this regime is designed: subject matter that incorporates more than trivial intellectual creation in the realm of communicative, expressive forms. Such forms must be distinguished from the content of expression, such as facts or ideas, and from functional aspects, all of which copyright does not target for protection. By demanding some modicum of creativity in forms protected by copyright, the originality requirement thus helps ensure, for example, that protected matter does not have its nature determined entirely by functional considerations that substantially override, if not obliterate, any sense of individual creativity in the resulting “work.”

In accordance with this view, a key point is that copyright law is not just about promoting newness or innovation per se, but instead is more specifically about promoting relevant forms of communicative expression, where, in this context, the word “forms” emphasizes the fact that copyright does not look to protect functional matter or content, such as facts or ideas. Requiring “creativity” means requiring more than trivial contribution in this domain of expressive forms. One can view attention to this, more specific aim of copyright as informing the Copyright Office’s approach to understanding § 102(a) and the statutory phrase “work of authorship”: as discussed earlier, in the Office’s view, for yoga sequences to qualify for copyright protection as, works of “choreographic authorship,” they must not only be “choreographic” but must reflect some creative or distinctively expressive aspect beyond those characteristic of “social dance steps and simple routines.”244 Likewise, the originality requirement and § 102(b) work

243 See William W. Fisher III, Recalibrating Originality, 54 HOUSE L. REV. 437, 461–65 (2016) (arguing that a higher originality threshold would protect works that add to cultural richness while freeing derivative works for “incorporation, appropriation, or recoding”).

244 Supra text accompanying note 228 (internal quotation marks omitted).
together to tailor copyright’s reach to the aim of promoting relevantly
creative expression.245

Because both the originality requirement and § 102(b) are meant to
prevent the extension of copyright to subject matter that is not in the domain
of relevantly creative expression, there should be little wonder that there is
substantial overlap in their application. There will be cases in which the
subject matter at issue is both excluded as functional under § 102(b) and
lacks the spark of creativity required for originality.246 In this zone of overlap
in outcomes, the value of redundancy between § 102(b) and the originality
requirement is mainly in providing alternative routes that can minimize
errors, a non-trivial benefit in an area where legal questions can seem opaque
or challenging for decision-makers.

Yet the overlap is not complete. There are cases where the creativity bar
of originality is, at least arguably, met but the relevant subject matter is still
excluded as functional. In Bikram’s Yoga, the Ninth Circuit illustrated this
phenomenon through examples of creative methods of brushing teeth,
mowing lawns, or drilling.247 And there are other cases where there are no
clear functionality problems under § 102(b), but nevertheless the claimed
subject matter fails the creativity test. This may happen either when after
deducting subject matter excluded by § 102(b) any remaining expressive
material is extremely thin248 or when, more generally, the relevant subject
matter does not fall within a § 102(b) category but includes no or little
creative expression of the nature that copyright is meant to protect.249

Under this understanding, § 102(b) and the originality requirement serve
a shared purpose even though there is only partial overlap between the two.
Both doctrines help ensure that copyright encompasses only expressive
forms for which it is designed. Section 102(b) achieves this by identifying
and excluding specific categories of subject matter outside this realm. The
originality requirement achieves this by focusing on a quality that cuts across

245 We thank Talha Syed for helping to sharpen this point.
Cir. 2004) (finding for purposes of a motion for a preliminary injunction that a short computer program
was ineligible for protection both because it failed to satisfy the originality requirement and because it
was functional due to its operation as a lockout code).
247 See Bikram’s Yoga, 803 F.3d at 1044 (presenting these examples).
248 Arguably this was the case in the seminal case of Feist Publications, Inc. v. Rural Telephone
Service Co., 499 U.S. 340, 362-364 (1991) (ruling that the standard selection, coordination, and
arrangement of unprotectable factual information in a white pages directory did not satisfy the
requirement of a modicum of creativity); see also Satava v. Lowry, 323 F.3d 805, 812 (9th Cir. 2003)
(finding that a combination of unprotectable elements in a glass in glass jellyfish sculpture fails the
originality requirement because the quantum of creativity “added in combining these standard and
stereotyped elements must be considered ‘trivial!’ ”).
249 See, e.g., Meshwerks, Inc. v. Toyota Motor Sales U.S.A., Inc., 528 F.3d 1258, 1265 (10th Cir.
2008) (ruling that barebones wire-frame computer models of cars fail the originality requirement because
they “reflect none of the decisions that can make depictions of things or facts in the world” protectable
by copyright).
different subject-matter types—its reflection of a modicum of creativity. Although there are some works for which the operation of the doctrines is complementary, with only one doctrine presenting a substantial barrier to copyrightability, the overall doctrinal structure is one of partial overlap in which, for nontrivial sets of works, both doctrines present meaningful hurdles, and the relatively minimalist creativity criterion functions as a stopgap that supports the § 102(b) exclusions, helping to dispose of copyrightability claims that might slip by the statutory exclusions but that nonetheless are not the type of subject matter that copyright is properly understood to cover. At least if one accepts the desirability of prioritizing avoidance of excessive copyright coverage, there can be value in this redundancy for works that hover around the borders of copyrightability and thus seem especially liable to slip by one of the individual doctrinal barriers even while being of dubious copyrightability under both.250

Hence, there seems a relatively easy case for value in retaining the originality requirement and § 102(b) exclusions as partially overlapping doctrines. Recognizing them as separate doctrines seems analytically useful, facilitating a care and precision in analysis that themselves can limit errors and promote the predictability and consistency of a legal regime’s operation. At the same time, permitting the doctrines’ overlap helps ensure exclusion from copyrightability in situations where the case for copyright protection is, in a sense, doubly weak.

Once we have the benefit of both the originality requirement and the § 102(b) exclusions, however, is there any work left for a distinct requirement of a “work of authorship”? We believe there is. The crucial point here is that functionality and the associated § 102(b) categories fail to capture all subject matter that should be categorically excluded from copyright. What of asana sequences in which no health or psychological benefit is asserted? What of sports moves, such as platform dives or gymnastics routines? Some of these may not be easily described as functional. Nor do they have a feature sometimes highlighted by courts as a basis for denying copyright protection to sports events: being “unscripted” and therefore authorless in the sense that the display does not follow a pattern determined by a well-defined creator.251 Should this mean, as one

---

250 To the extent considering a work’s creativity separately could help prevent a decision maker from failing to identify a subset of creative aspects that should not be refused copyrightability as a result of the § 102(b) exclusions, the originality requirement’s separate focusing of attention on creativity could also improve the operation of the copyright system in a way that, at least on occasion, favors would-be copyright holders. But for simplicity, we focus in the main text on how the partially redundant filters of § 102(a), § 102(b), and the originality requirement help prevent erroneous findings of copyrightability, rather than the extent the combination of these doctrines could help ensure that at least some aspects of certain works receive some copyright protection.

251 See Nat’l Basketball Ass’n v. Motorola, Inc., 105 F.3d 841, 846 (2d Cir. 1997) (“Sports events are not ‘authored’ in any common sense of the word.”).
commentator concludes, that such subject matter should generally be protectable by copyright? Our sense is that answer is “No.”

There are multiple reasons to believe that copyright exclusions, even for “original” subject matter, should extend beyond the functional and the non-creator-determined. In some areas of human ingenuity, there are sufficient incentives for creation even without exclusionary rights and thus no reason to incur the various costs of a proprietary regime. Consider the various first-mover and reputational advantages available to the innovator in the realm of sports moves or martial arts, where the innovations can provide a competitive advantage and a distinctiveness from the crowd that enables an early adopter win fame, fortune, or both. Similarly, consider the

---

252 See PATRY, supra note 52, § 4.21 (concluding that there is no bar for the protection of “certain sports like figure skating or synchronized swimming routines . . . [i]f the other elements of originality and fixation are present”).

253 See Samuelson supra note 225, at 55 (identifying as a central guideline for whether particular subject matter should be included in copyright the inquiry of whether “copyright or copyright-like protection is needed to induce investment in works of that kind.”).

254 First-mover or lead-time advantages are the advantages enjoyed by an innovator in a particular field by virtue of being the first to innovate especially when that person enjoys a period of time before others can catch up and emulate the innovation. See, e.g., F.M. SCHERER, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 384–85 (1973) (discussing lead-time advantages that can enable innovators to “enjoy freedom from competitive innovation for a sufficiently long time to recoup their original investment manyfold even without patent protection”); Stephen Breyer, The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs, 84 HARV. L. REV. 281, 300 (1970) (noting that, even at a time when the works of “many English writers” did not enjoy copyright protection in the United States, “lead time was important enough that [they] earned more from the sale of advance proofs to American publishers . . . than from the copyright royalties on their English sales”).

255 Reputational advantages are the positive effects enjoyed by an innovator due to being known as responsible for a particular innovation or series of innovations. See Glynn S. Lunney, Patent Law, the Federal Circuit, and the Supreme Court: A Quiet Revolution, 11 SUP. CT. ECON. REV. 1, 56–57 (2004) (discussing how “a reputation for innovation could become both a source of rents for, and an informal guarantee of, future innovation”); Rebeca Tushnet, Copy This Essay: How Fair Use Doctrine Harms Free Speech and How Copying Serves It, 114 YALE L.J. 535, 541 (2004) (observing that some creators “seek public accolades” or “want peer recognition”). Such positive effects could be “intrinsic” in the sense of the satisfaction enjoyed by the innovator as a result of the reputation irrespective of monetizing such reputation, or they could be pecuniary in the sense of leveraging the reputation into various forms of monetary income. See William M. Landes & Richard A. Posner, An Economic Analysis of Copyright Law, 18 J. LEGAL STUD. 325, 331 (1989) (discussing pecuniary and nonpecuniary advantages from the “self-advertisement and self-promotion” gains of publishing); Elizabeth L. Rosenblatt, A Theory of IP’s Negative Space, 34 COLUM. J. L. & ARTS 317, 343 (2011) (discussing pecuniary and nonpecuniary advantages that creators can gain from “recognition”). An example of reputational effects in sports is being known as the creator of a certain innovative element or even having the element named after the creator, see infra note 256—reputational devices that can generate for the innovator both intrinsic benefits and pecuniary income—for example, in the form of an increased demand for her coaching services.

256 Cf. ROBERT PATRICK MERGES & JOHN FITZGERALD DUFFY, PATENT LAW AND POLICY: CASES AND MATERIALS 203 (6th ed. 2013) (noting that Dick Fosbury used his early advantage in “perfect[ing] the back-first high-jump technique now known as the ‘Fosbury Flop’ . . . to set an Olympic record and win the gold medal in the 1968 Olympics”).
alternative sources of motivation and compensation\textsuperscript{257} available to the organizers of a parade,\textsuperscript{258} the creator of a culinary dish,\textsuperscript{259} or the designer of a new perfume fragrance.\textsuperscript{260}

More fundamentally, in some areas, proprietary rights may be corrosive of the very value of the human activities and relationships at stake. Sport is a good example. As one court put it, “\textit{e}ven where athletic preparation most resembles authorship—figure skating, gymnastics, and, some would uncharitably say, professional wrestling—a performer who conceives and executes a particularly graceful and difficult—or, in the case of wrestling, seemingly painful—acrobatic feat cannot copyright it without impairing the underlying competition in the future. A claim of being the only athlete to perform a feat doesn’t mean much if no one else is allowed to try.”\textsuperscript{261}

Hence, by providing grounds for defining what lies within the bounds of the copyrightable, § 102(a) is itself a means for rough categorical exclusions distinct from those prescribed by Congress in § 102(b), such as the exclusion based on functionality.\textsuperscript{262} Notably, § 102(a) embodies a distinctive congressional strategy for limiting copyright protection to appropriate matter—namely and in part, listing categories of matter that characteristically falls within copyright’s proper realm. More specifically, § 102(a) creates a division of labor between Congress and the courts in the deployment of this additional filter. Congress assumes the primary responsibility for designating areas of subject matter deemed suitable and

\textsuperscript{257} See Christopher Buccafusco, \textit{Making Sense of Intellectual Property Law}, 97 CORNELL L. REV. 501, 546–48 (2012) (arguing that, although various innovations in fields such as movement, tactile or ergonomic products, and culinary dishes can have nonfunctional elements, it may be desirable to deny them copyright protection because of the existence of sufficient independent sources of motivation and incentives).

\textsuperscript{258} See Prod. Contractors, Inc. v. WGN Cont’l Broad. Co., 622 F. Supp. 1500, 1504 (N.D. Ill. 1985) (concluding that a parade, as opposed to an audiovisual depiction of the parade, is “not a work of authorship entitled to copyright protection”).

\textsuperscript{259} Courts routinely deny copyright protection to new culinary creations. They often do it by finding a new recipe to be functional. \textit{See}, e.g., Publ’ns Int’l, Ltd. v. Meredith Corp., 88 F.3d 473, 482 (7th Cir. 1996) (denying copyright protection to yogurt recipes because they were found to be functional under § 102(b)). Some courts, however, have relied on a different reason that a new culinary creation is not encompassed by any § 102(a) subject matter category, thereby implying that even possibly nonfunctional elements of culinary creations are not entitled to copyright protection. \textit{See}, e.g., Lorenzana v. S. Am. Rests. Corp., 799 F.3d 31, 34 (1st Cir. 2015) (denying copyright protection to a chicken sandwich because it was found not to fit “any of the eligible categories”).

\textsuperscript{260} See Charles Cronin, \textit{Genius in a Bottle: Perfume, Copyright, and Human Perception}, 56 J. COPYRIGHT SOC’Y, U.S.A 427, 471 (2009) (positing that, “like the products of the closely related apparel industry, perfume is a ‘solidarity good’” and thus can be produced “profitably and innovatively” despite only limited patent, trademark, and trade secret protection).

\textsuperscript{261} Nat’l Basketball Ass’n v. Motorola, Inc., 105 F.3d 841, 846 (2d Cir. 1997).

\textsuperscript{262} See NIMMER & NIMMER, supra note 103, § 2.03[A] (observing that “the concept ‘works of authorship’ is intentionally left vague under the Act” and that “courts are thereby permitted, but not required, to recognize as protectable types of works not expressly included in the category enumeration set forth in the statute”).
therefore eligible for copyright protection through explicitly specified subject matter subcategories.\textsuperscript{263} Courts are given secondary responsibility that they exercise in two ways. First, courts decide through interpretation whether borderline cases fall within the categories enumerated by Congress, their reserved discretion to make such determinations providing for flexibility in calibration of copyright’s subject-matter coverage.\textsuperscript{264} Second, courts are left to make judgments on the suitability for copyright protection of forms of human creation outside the enumerated subject-matter categories: courts make such judgments in deciding whether such matter properly falls within the residual area that lies outside the enumerated categories but nonetheless inside the overarching category of “works of authorship.”\textsuperscript{265} These judgments can proceed by analogy to existing recognized categories, thus giving the statutory list of explicit inclusions another significant purpose.\textsuperscript{266} Of course, one can also give §102(a)’s subject matter requirement significant, independent meaning by adopting the Copyright Office’s strict and static view of subcategories of copyrightable subject matter, thus limiting the court’s responsibility to managing and interpreting enumerated subject matter categories.\textsuperscript{267} Whether one adopts this restrictive view or a broader, more flexible understanding of §102(a), however, the work-of-authorship requirement can and should function as yet another preliminary filter to enable the exclusion of subject matter that is inappropriate for copyright.

In sum, originality, §102(a) subject matter requirements, and §102(b)

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{263} E.g., 17 U.S.C. § 102(a) (2012) (specifying eight categories of works of authorship); see also Golan v. Holder, 565 U.S. 302, 332 (2012) (“Congress recurrently adjusts copyright law to protect categories of works once outside the law’s compass.”).
\item \textsuperscript{264} See Reese supra note 225, at 1513 (describing how Congress can attain judicial and administrative flexibility in calibrating the subject matter coverage of copyright by defining enumerated categories in broad, conceptual terms and relying on courts and the Copyright Office to interpret and apply them).
\item \textsuperscript{265} See Nimmer Nimmer, supra note 103, § 2.03[A] (noting the ability of courts to recognize additional types of works beyond those statutorily enumerated).
\item \textsuperscript{266} Cf. Gideon Parchomovsky & Alex Stein, Catalogs, 115 COLUM. L. REV. 165, 168 (2015) (describing use of “catalogs” to indicate the scope of a legal super-category through “a specific enumeration” of illustrative included matter along with an allowance for further inclusions according to “the family-resemblance principle”). It seems that the legislative history of the 1976 Act reflects a similar view of the division of labor with respect to categories of copyrightable subject matter and the constraining role within it of enumerated categories. See H.R. REP. No. 94-1476, at 51 (1976) (“Authors are continually finding new ways of expressing themselves, but it is impossible to foresee the forms that these new expressive methods will take. The bill does not intend either to freeze the scope of copyrightable subject matter at the present stage of communications technology or to allow unlimited expansion into areas completely outside the present congressional intent. Section 102 implies neither that that subject matter is unlimited nor that new forms of expression within that general area of subject matter would necessarily be unprotected.”).
\item \textsuperscript{267} See supra text accompanying notes 221–23 (describing the Copyright Office’s strict approach); Reese supra note 225, at 1499 (arguing that a reformed statute “should not extend protection to an open-ended and undefined category of works”).
\end{enumerate}
\end{footnotesize}
exclusions overlap in their coverage and share a common purpose. There is subject matter that is functional, fails to meet the minimum threshold of originality’s creativity aspect, and is also outside § 102(a)’s enumerated or residual categories of “works of authorship.” But these doctrines have nonoverlapping portions of their scopes and significant differences in their means and operational logic. The originality requirement demands case-by-case analysis of whether a particular work exhibits a minimal degree of creativity in the realm of expressive, communicative forms. Section 102(b)’s functionality exclusion screens out works or aspects of works, no matter how creative, in essence on the ground that their utilitarian nature renders them inapt for copyright protection. Section 102(a) demands that the courts affirmatively find that subject matter is properly included within the realm of copyrightable subject matter based on its location within either a specifically enumerated or a residual category of “works of authorship.” Section 102(a) thereby serves as a general-purpose filter that allows courts to deny entry to the realm of copyright for entire subfields of human creation where, on balance, proprietary rights are likely to do more harm than good. One can make the case that where these filters are partially redundant in their blocking of copyright’s availability, the relevant subject matter is triply suspect as a candidate for copyright protection from a social-welfare perspective. In this sense, the layering of § 102(a) and § 102(b) subject-matter eligibility restrictions on top of a backstopping requirement of originality exemplifies how partial redundancy in legal design can be a social good rather than a bane.

D. Reproduction and Derivative-Work Rights

The last example of redundancy in copyright law is a cautionary one, illustrating how redundancy can go astray. This example relates to the relationship between the right of reproduction and the right of preparing derivative works. These are two of a copyright owner’s exclusive entitlements enumerated in § 106 of the Copyright Act. Mainly for historical reasons, there is substantial overlap between the two. The


269 See Bikram’s Yoga Coll. of India, L.P. v. Evolation Yoga, L.L.C., 803 F.3d 1032, 1044 (9th Cir. 2015) (observing that, without § 102(b)’s functionality exclusion, “one might obtain monopoly rights over [] functional physical sequences by describing them in a tangible medium of expression and labeling them choreographic works”).

270 See Nat’l Basketball Ass’n v. Motorola, Inc., 105 F.3d 841, 846 (2d Cir. 1997) (rejecting copyright protection of basketball games because such “athletic events” do not fall within a category of “works of authorship” explicitly listed by § 102(a) and they are “neither similar nor analogous to any of the listed categories”).

reproduction right is copyright’s core entitlement of making copies—fixing the work in a tangible medium of expression. The derivative-work right of § 106(2) confers on the owner a broad exclusive right to make secondary works based on the original, including, for example, translations, musical arrangements, dramatizations, motion picture adaptations, abridgments, and sound recordings.

Many scholars have argued that the derivative-work right should be restricted in various ways or even abolished, or that a circumscribed reading of it is consistent with Congress’s intent in recognizing the right statutorily in 1976. Courts, however, have repeatedly read the derivative-work right very broadly, pointing to statutory language defining the term “derivative work” as encompassing any “form in which a work may be recast, transformed, or adapted.”

Meanwhile, courts have also read the § 106(1) right of reproduction broadly. In modern copyright law, this right extends well beyond verbatim reproduction or even verbatim reproduction with minimal changes. It encompasses any copying that generates a result substantially similar to the original, where substantial similarity is capaciously understood to include a

---

272 See id. § 106(1) (creating an exclusive right to reproduce the work “in copies or phonorecords”).
273 See id. § 101 (defining “copies”).
274 See id. § 106(2) (creating an exclusive right to “prepare derivative works based upon the copyrighted work”); id. § 101 (defining a “derivative work”).
275 See, e.g., Derek E. Bambauer, Faulty Math: The Economics of Legalizing the Grey Album, 59 ALA. L. REV. 345, 391 (2008) (asserting four key benefits to abolishing the derivative works right); Christina Bohannon, Taming the Derivative Works Right: A Modest Proposal for Reducing Overbreadth and Vagueness in Copyright, 12 VAND. J. ENT. & TECH. L. 669, 692–94 (2010) (arguing that the Copyright Act, including the derivative works right, is overbroad); Tyler T. Ochoa, Copyright, Derivative Works and Fixation: Is Galoob a Mirage, or Does the Form(Gen) of the Alleged Derivative Work Matter?, 20 SANTA CLARA COMPUTER & HIGH TECH. L.J. 991, 1020 (2004) (suggesting that the right to prepare derivative works is “not independent of the other four exclusive rights, but is infringed only in conjunction with at least one of the other four exclusive rights”); Stewart E. Sterk, Rhetoric and Reality in Copyright Law, 94 MICH. L. REV. 1197, 1215–17 (1996) (asserting that arguments and justifications supporting “giving an author exclusive rights in derivative works” are unpersuasive and such protection ceases to generate “significant incentives for creative activity”); Naomi Abe Voegtli, Rethinking Derivative Rights, 63 BROOK. L. REV. 1213, 1267–68 (1997) (proposing to not consider most transformative works as derivative works); Tim Wu, Tolerated Use, 31 COLUM. J.L. & ARTS 617, 631 (2008) (arguing that one such approach to solving “some of the problems of tolerated use” is to “suggest that a product that is a true complement to the original work . . . simply should not be considered an adaptation of the work under section 106(2)”).
276 See Pamela Samuelson, The Quest for a Sound Conception of Copyright’s Derivative Work Right, 101 GEO. L.J. 1505, 1511 (2013) (contending that “the exclusive right to prepare derivative works is narrower in scope and more bounded than commentators have often feared”).
277 17 U.S.C. § 101 (2012); see Samuelson, supra note 276, at 1509–10 (asserting that the derivative work right “has been highly problematic in some cases in which courts have interpreted [the right] broadly”). The most prominent scholarly support for a broad approach to the derivative-work right appears in the treatise by Paul Goldstein, 2 PAUL GOLDSTEIN, GOLDSTEIN ON COPYRIGHT § 7.0 (3d ed. 2011).
wide—although not unlimited—range of similarities.278

The broad scope accorded to each of the two rights is the source of their overlap. Is a sequel for a novel that uses many of its characters and some of its plot lines a copy, a derivative work, or perhaps both? The usual answer by courts is that the specific classification as a copy or a derivative work does not matter.279 The violation of either right is subject to the same infringement test, defenses, and remedies.280 Generally speaking, therefore, there seems little need to try to define a clear dividing line between the two rights and thus to require the wasting of energy classifying individual cases of infringement as falling on one or the other side of the line.

Indeed, some have pushed this logic further, arguing that in light of the reproduction right and other § 106 entitlements, the derivative-work right is “completely superfluous.”281 It is generally recognized, however, that various doctrinal technicalities turn on the issue of a work classified as a reproduction or a derivative work.282 Sometimes the cause is a statutory distinction. Various rights of authors to terminate copyright transfers under § 203(a)283 and § 304(c)–(d),284 for example, are subject to limitation based on the reliance interests of those who made authorized derivative works prior to the termination.285 Such users enjoy a privilege to continue exploiting these derivative works notwithstanding the termination.286 A related privilege to exploit prior non-infringing derivative works exists under

---


279 Usually this assumption is not even made explicitly but rather is implied in a casual and indifferent approach by courts to the question of the reproduction/derivative-work borderline. A representative example of this casual approach is Castle Rock Entm’t, Inc. v. Carol Publ’g Group, Inc., 150 F.3d 132 (2d Cir. 1998). In this case the court held that a trivia game that derived its questions and answers from the details of a television program infringed on the television program’s copyrights. Id. at 138–39. The plaintiff’s argument was that the game was a derivative work, but the court’s opinion demonstrates a conspicuous indifference to the question of which entitlement is infringed and even refers repeatedly to the game as a work that “copies” the program. Id. at 138–41.


281 See NIMMER & NIMMER, supra note 103, § 8.09[A][1] (presenting the argument that the derivative-work right is superfluous but then noting “some exotic situations in which the adaptation right may take on substantive significance”).

282 Compare id. § 3.04 (addressing the scope of protections for derivative work), with id. § 8.02 (addressing the same for reproductions).

283 17 U.S.C. § 203(a) (2012) (providing a right to terminate a transfer or license under specified conditions, including the lapse of thirty-five to forty years since the transfer or license).

284 Id. § 304(c)–(d) (2012) (providing right to terminate certain transfers or licenses of renewal copyright for certain works).

285 See, e.g., Stephen McJohn, The Case of the Missing Case: Stewart v. Abend and Fair Use, 53 IDEA 323, 347 (2013) (noting that § 203’s “termination provisions . . . allow continued use of derivative works prepared under the terms of the grant (in other words, in reliance on the grant”)).

286 See id. §§ 203(b)(1), 304(c)(6)(A) (providing conditions for continued use of a derivative work after termination of a transfer or license); cf. id. § 304(a)(4)(A) (providing conditions for continued use of a derivative work under a class of situations involving a renewed or extended copyright term).
§ 104A, \textsuperscript{287} which restores copyright protection to foreign works that fell into the public domain due to non-compliance with formality requirements under former U.S. law.\textsuperscript{288} In both of these cases, the privilege applies only to derivative works and not to reproductions.

In other cases, the need to distinguish between a reproduction and a derivative work arises from a more fundamental feature of the copyright owner’s rights. For example, consider a case where I purchase a lawful copy of Rembrandt’s \textit{The Night Watch}, hypothetically under copyright, and I add little green alien figures in the background. I did not reproduce the original work because I did not create a copy: I did not generate a new tangible fixation of the original. But as long as what I created is a derivative work, I potentially infringed the derivative-work right, which does not require the making of a copy.\textsuperscript{289} In such cases, it becomes important to recognize the partial independent coverage of the right. Courts often struggle both in defining the copyright owners’ separate rights and in classifying arguably infringing activities accordingly.\textsuperscript{290} Arguably, however, these cases are sufficiently manageable and containable to justify the otherwise careless approach to the only partially overlapping nature of reproduction and derivative-work rights.

Another set of cases is more worrisome. Here the careless and undisciplined approach to the overlap between reproduction and derivative-work rights has brought about confused and even absurd results—alarmingly detached from background policies. Consider the case of \textit{Warner Bros. Entertainment Inc. v. RDR Books}.\textsuperscript{291} In this case, Steven Vander Ark, a fan of the \textit{Harry Potter} novels by J.K. Rowling, developed and operated a website that functioned as a lexicon on all things Harry Potter—your go-to

\textsuperscript{287} Id. § 104A(d)(3) (allowing for continued use of certain derivative works “if the reliance party pays to the owner of the restored copyright reasonable compensation” for otherwise infringing conduct).

\textsuperscript{288} See id. § 104A(a)(1)(A) (providing for copyright in “restored works”).

\textsuperscript{289} See Lee v. A.R.T. Co., 125 F.3d 580, 581 (7th Cir. 1997) (reasoning that “§ 106(2) creates a separate exclusive right” distinct from both the rights of reproduction and of distribution of copies). A related but distinct issue is whether a derivative work has to be fixed at all in order to infringe, as in the case of an unfixed live performance derivative of a copyrighted play. \textit{Compare} Lewis Galoob Toys, Inc. v. Nintendo of Am., Inc., 964 F.2d 965, 968 (9th Cir. 1992) (observing that “[a] derivative work must be fixed to be protected under the Act, . . . but not to infringe” (citation omitted)), \textit{with Patry}, supra note 52, § 12:14 (arguing that copyright “may be infringed by an unauthorized unfixed derivative work”), \textit{with Nimmer & Nimmer, supra} note 103, §8.09 [A][2] (explaining that the court’s pronouncement in \textit{Galoob} that an infringing derivative work may be unfixed is dictum and arguing that “fixation should be required to infringe the adaptation right”).

\textsuperscript{290} See A.R.T. Co., 125 F.3d at 581 (determining whether reprinting an artist’s work on tiles counts as a derivative work); Mirage Editions, Inc. v. Albuquerque A.R.T. Co., 856 F.2d 1341, 1343 (9th Cir. 1988) (reasoning that art reproduction is not necessary for finding a derivative work); Peter Lettersee & Assocs., Inc. v. World Inst. of Scientology Enters., Int’l, 533 F.3d 1287, 1299–1300 (11th Cir. 2008) (determining whether the use of checklists and drills from a religious book in a course is a derivative work of the original book).

\textsuperscript{291} 575 F. Supp. 2d 513 (S.D.N.Y. 2008).
place if you experience a burning urge to read, for example, an encyclopedia of spells and potions. For years after the creation of this website, Vander Ark received praise and even special treatment from J.K. Rowling and other stakeholders in the Harry Potter empire until he agreed to publish the Lexicon in print through defendant RDR Books. Vander Ark had previously worried about publishing the Lexicon in print, particularly after Rowling had indicated her own interest in writing a Harry Potter encyclopedia, but RDR Books had reassured him that the print publication would be legal and agreed to “defend and indemnify [him] in the event of any lawsuits.” Within a couple months of Vander Ark’s publishing agreement with RDR Books, the copyright holders filed suit.

Sensibly, the district court found that the Lexicon was not a derivative of the Harry Potter copyrighted work. The court reasoned that not every work inspired by another, or based on another in some loose sense, is a derivative in the copyright sense. More specifically, the court concluded the Lexicon—a collection of alphabetically organized entries about all characters, places, creatures, spells, and potions in the Harry Potter universe—was too different from the work of fiction known as Harry Potter to be considered a transformed or adapted version of it. In the court’s terms, the Lexicon was so transformative that it “no longer ‘represent[ed] [the] original work[s] of authorship’” and thus did not infringe the derivative-work right. This analysis seems plausible. In the very same opinion, however, the court held the Lexicon to be a copy of Harry Potter and therefore in violation of the reproduction right.

This is, of course, a perverse outcome. It could be reached only by reasoning that had become unmoored from the basic purpose of the doctrine in this area. Defending a conclusion that a particular use is not a derivative work but is an infringing reproduction requires some conceptualization of each of the rights and the relationship between them. But the primary conceptualization that comes to mind points in the opposite direction from

292 Id. at 520.
293 See id. at 521–22 (describing “positive feedback, including from Rowling and her publishers” before Vander Ark agreed to publish the Lexicon).
294 Id. at 522.
295 Id. at 524.
296 Id. at 539.
297 Id. at 538.
298 Id. at 520, 539.
299 Id. at 539 (alteration in original).
300 Id. at 534–38; see Samuelson, supra note 276, at 1541–42 (praising the court’s conclusion that the Lexicon was too different to be a derivative of the Harry Potter books). But any basis for praise here is severely, if not entirely, undercut by the court’s accompanying finding that the Lexicon was an infringing reproduction.
the court’s logic and conclusion.

An attractive candidate for explaining the relationship between the reproduction and derivative-work right is as follows. The two rights occupy different but partly overlapping spaces on a continuum of similarity to expressive elements of the original work.302 For simplicity, this continuum is represented as extending along a single horizontal direction in Figure 1 below, with the distinct scopes of the reproduction and derivative-work right indicated by overlapping triangles. In reality, however, the continuum of similarity extends along multiple dimensions reflecting medium of expression as well as content. As one commentator puts it, a derivative work lies in the “gray zone” on “the continuum between an exact reproduction of protected property[] and the creation of an original work.”303 A duplication of the text of a copyrighted novel with only minor changes to certain phrases is reproduction, but its extreme closeness to the original probably precludes it from being a derivative work. 304 A motion picture that serves as a prequel to the story told in a copyrighted novel and therefore uses some of its characters and incorporates several of its other expressive elements is not a reproduction, but it may be a derivative work.305 An annotated selection of a few of the novel’s chapters may be both a reproduction and a derivative work.306

---

302 Cf. id. at 807 (observing that “[t]he distinction between derivation and (other forms of) reproduction is that a nonderivative form of copying takes the expression of the primary work, while a derivation may take but it must also transfer and transform what makes the primary work original” (emphasis in original)); id. at 845 (showing a diagram presenting reproduction and derivative work as occupying different parts of a continuum of concrete expression taken from a primary work).

303 Timothy Everett Nielander, The Mighty Morphin Ninja Mallard: The Standard for Analysis of Derivative Work Infringement in the Digital Age, 4 TEX. WESLEYAN L. REV. 1, 2 (1997). The term “an original work” does not refer here to meeting copyright’s originality threshold, but to a work that is too different to be a derivative of another work from which it borrows. No doubt, a work may be a derivative work and satisfy the originality requirement.

304 See Gervais, supra note 301, at 788, 801–02 (explaining that the definition of a derivative work is not clear and that, although statutes provide examples of named derivatives, they do not provide an exhaustive list).

305 See id. (explaining how a “derivative work transforms or recasts something protectable in the primary work . . . by adding or transforming it”).

306 See id. (providing examples of derivative works and that “some additional work by the derivative user” is necessary).
Figure 1: Conceptualization of Overlap Between the Reproduction and Derivative-Work Right Along a Horizontal Axis of Similarity to the Original

The point of this conceptualization is not to draw a crisp line between the two categories, but instead to convey in a disciplined and coherent way the relationship between them as it relates to underlying policy. As the level of similarity between elements derived from a copyrighted work becomes increasingly remote from verbatim reproduction, the ratio of the social value added by copyright coverage of those elements and the social cost of such coverage can be expected to drop. This is so because, generally speaking, control of remote-similarity uses can be expected to be less significant for ensuring a robust level of incentives for primary authors, but can be expected to impose undiminishing or even increasing social costs through taxation or discouragement of secondary uses that likely incorporate original content, perhaps even more so as the distance from the primary work increases. As the value-to-cost ratio of copyright coverage drops, so does the justification for the extension of exclusive rights. This is all the more true from the point of view of normative outlooks that favor a robust freedom for secondary uses and creative re-appropriation beyond what is dictated by the

---


308 Cf. Louis Kaplow, The Patent-Antitrust Intersection: A Reappraisal, 97 Harv. L. Rev. 1813, 1829–30 (1984) (describing a way of assessing the relative desirability of different legal policies by examining “the ratio between the reward the patentee receives when permitted to use a particular restrictive practice and the monopoly loss that results from such exploitation of the patent”).
strict calculus of economic efficiency.\textsuperscript{309}

There is a lively debate about whether the current, broad derivative-work right can be justified under any normative framework, with most commentators answering in the negative.\textsuperscript{310} It is quite possible that the preferable alternative is an overall reduced scope of copyright achieved by eliminating or shrinking the derivative-work right and also obviating any backdoor derivative-work right by more tightly circumscribing the reproduction right. For our purposes, it is enough to point out that, at least if derivative-work right or its analog is not entirely eliminated, a proper understanding of the relationship between the reproduction and derivative-work right has important implications. In accordance with the conceptualization proffered here, a derivative-work right—and in particular that portion of such a right covering secondary uses sufficiently different from the original that they cannot be plausibly seen as reproduction—rests on a weaker and more precarious justification than the core of copyright’s reproduction right.\textsuperscript{311} It follows that as we advance on the continuum of similarity from the exclusive area of reproduction to the zone overlapping with derivative works, and then to that associated exclusively with derivative works, the justification for copyright thins. Once we proceed along this continuum to uses that do not incorporate enough of the expressive material of the original to count even as a derivative work, the justification for copyright protection runs out altogether.

Implications of this conceptual picture are twofold. First, perverse logic such as that of the \textit{RDR Books} decision should be avoided. If a court decides that a secondary work does not incorporate enough of a primary work’s authorship to count as a derivative, it should not find that the secondary work is a reproduction. What is too different to be a derivative work is \textit{a fortiori} too different to be a reproduction. Second, as we proceed along the continuum of similarity away from exact reproduction, courts generally should be cognizant of the corresponding weakening justification for copyright protection and therefore warier of imposing liability or heavy copyright remedies, even when an allegedly infringing work is found to be a reproduction or derivative. Practically this means that as they enter the derivative-work zone, and particularly its outer parameter, courts should apply other constraining doctrines, such as fair use, \textit{scènes à faire}, and the idea-expression dichotomy with greater care and frequency to prevent

\textsuperscript{309} Bracha & Syed, \textit{supra} note 195, at 269–74 (explaining how self-determination theory provides stronger support than efficiency for a robust freedom to create works that creatively re-appropriate material from preexisting works).

\textsuperscript{310} See \textit{supra} notes 275–60 and accompanying text (explaining the definition of a derivative work and right, the relation of a derivative work to an original, and the benefits of eliminating the derivative-work right).

\textsuperscript{311} See Bambauer, \textit{supra} note 275, at 391 (stating that the derivative-work right is weak and should be abolished).
copyright law from overreaching its social purpose.\textsuperscript{312}

The foregoing analysis suggests that the difficulty with the reproduction and derivative-work right interface is not the partial overlap as such. Redundancy does not by itself cause perversity such as that seen in \textit{RDR Books}. Difficulty can arise, however, from a failure to ground decision-making in a coherent and, from a policy perspective, plausible conceptualization of the relationship between overlapping doctrines. Once overlapping doctrines are unmoored from their underlying purposes and left to float as formal shells, all sorts of strange and undesirable outcomes can follow. The trouble, in other words, is not redundancy per se, but redundancy whose structure is not well-designed or implemented in light of the substantive policies of the relevant legal field.

\textbf{CONCLUSION}

As a principle of design for legal doctrine, redundancy can help the law secure core interests, allow for case-by-case flexibility and long-term evolution, and do so in ways that—relative to alternatives—save on legal drafting and administration costs. But of course, redundancy can be poorly implemented or overdone, a fact that at least partly explains the common rhetorical dominance of anti-redundancy principles. Greater appreciation of the potential positive and negative aspects of redundancy in the context of underlying policies can helpfully inform evaluation of the complex array of legal doctrines that regulate copyrightability and liability for copyright infringement. Further, better attention to the subtleties of redundancy, including the various potential forms of partial redundancy, can help ensure that redundancy does not go awry.

Our examples of doctrinal redundancy in copyright law help illustrate these general principles, as well as additional aspects of how redundancy can and does operate in practice. Three of these examples illustrate potential rationales for keeping redundant limitations on copyright protection. But we should emphasize that redundancy can also work in tandem with robust copyright protection, as it has historically through the emergence of derivative-works rights on top of expansive reproduction rights.\textsuperscript{313} A fuller discussion of the main takeaways from our copyright examples follows.

\textsuperscript{312} See Gervais, \textit{supra} note 301, at 809–10 (stating courts need to evaluate additional factors and that the “idea-expression dichotomy and … the \textit{scènes à faire} doctrine[,] are other good candidates for a differential application to derivative uses”); Samuelson, \textit{supra} note 276, at 1554–58 (discussing cases that distinguish derivative works from infringement).

\textsuperscript{313} See \textit{supra} text accompanying notes 260–63 (explaining that the derivative-work right and the right of reproduction overlap but there are cases in which it is necessary to distinguish between the two). This case of pro-copyright redundancy does provide us with an example of how redundancy can go wrong, but, in our account here, that example is fundamentally one of misapplication following from inattention to underlying purpose, not an example of the inadvisability of such pro-copyright redundancy per se. See \textit{supra} text accompanying notes 276–81.
The partial redundancy between copyright’s fair use doctrine and § 102(b)’s exclusion of functional subject matter illustrates a form of rule-standard redundancy in which a more rule-like doctrine—here, exclusion of functional subject matter from copyrightability—at least partly overlays a more general, jack-of-all-trades standard—here, the exemption of fair uses from copyright liability. Although the question of what is considered functional for copyright purposes can at times seem standard-like, particularly when applying doctrines such as the merger doctrine, the determination of functionality has a relatively narrow focus in comparison to the relatively totality-of-circumstances nature of fair use analysis. Consequently, § 102(b)’s requirement of nonfunctionality can channel certain functional material away from copyright protection more readily and predictably than would be likely under even a very robust version of the fair use exemption alone. Further, part of the reason that the nonfunctionality requirement can achieve this is procedural and institutional: the less factually intensive nature of standard functionality inquiries makes them better tools for decisive action by trial judges in response to motions for early dismissal or summary judgment.

At the same time, retaining a robust fair use backstop can also add substantial value, particularly at copyright’s “frontier,” where courts are seeking to apply old nonfunctionality doctrines in new technological and business environments. In such areas, courts might not always apply the nonfunctionality requirement properly. The courts’ struggles with respect to questions of interoperability in the Oracle case and its analogs illustrate this point. In other situations, entrenched doctrinal features may frustrate reaching desirable results under new and unpredicted technological or economic circumstances, as in the Sega case. In both of these sets of circumstances in which courts are especially error-prone, there seems to be good reason to deploy a strong backstopping principle, such as the open-textured doctrine of fair use. The combination of a rule-like overlay and a flexible backstopping standard can thereby enable society to substantially enjoy the advantages of a rule-like norm in a plurality of core cases while also retaining a safety net to prevent frustration of underlying policies.

314 See supra text accompanying notes 129–36 (discussing the Sega case).
315 See Samuelson, supra note 5, at 427–28 (“Courts in numerous merger cases have taken into account whether the alternative expressions were inferior in some way, such as because they were less efficient, impractical, unreasonable, illogical, or contrary to industry expectations.” (footnotes omitted)).
316 See supra text accompanying notes 107–09 (explaining that outcomes under the fair use doctrine are hard to predict because of its complex inquiry, as compared to outcomes under § 102(b), which it is more straightforward).
317 See supra text accompanying notes 109–11 (discussing procedural considerations relating to the nonfunctionality requirement).
318 See supra text accompanying notes 81–84 (discussing how the court in the Oracle case handled the question of interoperability).
319 See supra text accompanying notes 129–36 (discussing the Sega case).
through changed circumstances or conscious circumvention.

Of course, under certain circumstances, a rule-plus-standard combination could lead to society predominantly suffering from the disadvantages of both a rule and a standard. But this unhappy result can be rendered less likely by applying two design principles. First, the reinforcing aspects of the rule and standard should be such that they help prevent errors in a direction that is especially disfavored—for example, because of background presumptions in favor of free competition or speech. Second, to the extent there is substantial cause for concern about, say, the unpredictability of the standard with respect to reasonably well-established forms of activity, the standard might be confined to a relatively secondary role, predominantly involving extreme or novel circumstances.

Copyright’s doctrines of fair use and improper appropriation exemplify a different form of redundant structure: one involving two standards having overlapping concerns, but fundamentally involving distinctly structured inquiries. Although improper appropriation doctrine’s substantial similarity inquiry is about as standard-like as one can imagine, this inquiry does focus sharply on a comparison of the expressive content of two works—a comparison that tends to form only one aspect of the fair use doctrine’s more totality-of-the-circumstances inquiry. Here, the more limited focus of the improper appropriation standard can, for a certain class of situations, render its relation to fair use effectively like that of a rule laid over a more general, backstopping standard. In situations where lack of substantial similarity may be properly found based on a streamlined set of readily accessible and perhaps even relatively undisputed facts, the improper appropriation doctrine can enable disposition of a case more quickly, cheaply, and predictably than a fair use inquiry.

As noted earlier, the nonfunctionality doctrine can similarly enable such quicker, cheaper, and predictable dispositions in a distinct subset of cases. In both cases—that of the improper appropriation doctrine and that of the nonfunctionality doctrine—these advantages extend beyond mere improvements in the administrability of legal processes. In the copyright context, easier and more predictable processes for solving disputes and dismissing meritless claims promotes underlying substantive policies by

---

320 See supra text accompanying note 206 (suggesting that the overlap between fair use doctrine and the requirement of improper appropriation is one way to obtain more desirable results); Golden, supra note 6, at 707 (“[W]here law is focused on one dominant interest and we are confident that we will want the law to retain this focus over time, redundancy [to protect that interest] might well be a good social strategy.”).

321 See Golden, supra note 6, at 709 (discussing “contract law’s unconscionability doctrine”).

322 See supra text accompanying notes 154–74 (comparing the doctrines of fair use and improper appropriation and illustrating their similarities and differences in the context of Warner Bros. Inc. v. Am. Broad. Cos., 720 F.2d 231 (2d Cir. 1983)).

323 See supra text accompanying notes 157–56 (describing the improper appropriation doctrine as having a “narrow focus” and the fair use doctrine as being “multi-focal”).
helping prevent undue chilling of speech liable to attract accusations of copyright infringement. Process improvements thus make it more likely that relevant parties will actively enjoy the use privileges given to them.

Once again, however, there is a potential counterbalancing concern. The layering of different legal hurdles to copyright liability can generate more false negatives in the form of socially undesirable determinations of no liability, even as the layering decreases false positives in the form of socially undesirable determinations of liability. At least to some degree, this concern might be viewed as sharper in the case of layered standards than in the case of a rule overlaid on a standard because the layered-standard structure might provide two broad bases for uncertainty without the sort of “safe harbor” or “sure shipwreck” predictability possible within the core region encompassed by a rule.324 Within the context of copyright, however, we think the critical value attached to background free speech interests can justify a layering of bulwarks for those interests, despite associated costs.325 Indeed, the core nature of the free speech interest in democratic society may demand such an arrangement. In this context, well-designed redundancy is a doctrinal strategy for safeguarding “higher-order” interests even at some possible detriment to other interests or goals.

Copyrightability requirements of originality, nonfunctionality, and qualification as a work of authorship illustrate a third form of doctrinal structure: one of multiple overlapping filters, each imperfect, that through collective action serve a shared purpose of providing proper boundaries for the scope of a legal field. In a significant sense, the originality, work-of-authorship, and nonfunctionality requirements break the question of copyrightability down into separately identifiable, but not wholly distinct parts, each with a different focus and structure for inquiry.

The work-of-authorship and nonfunctionality requirements both operate as relatively coarse-grained filters, tending to function at the level of types appropriate or inappropriate for copyright protection.326 The exclusion of functional subject matter in § 102(b) tends to have a unidimensional function and focus: channeling a specific kind of subject matter for which the features of copyright seem a poor fit away from this regime and in the direction of

324 See Susan C. Morse, Safe Harbors, Sure Shipwrecks, 49 U.C. DAVIS L. REV. 1385, 1387–88 (2016) (describing standards as “provid[ing] general considerations . . . [that] leave the determination of legal results to future decision makers,” while describing the ability of “safe harbor[s]” and “sure shipwreck[s]” to guarantee that particular facts will result in particular legal outcomes).
325 See supra text accompanying notes 200-11 (discussing how the doctrinal layering of the fair use and improper appropriation doctrines helps protect free speech interests).
326 See supra text accompanying notes 231–34, 237–41 (discussing the role of § 102(a) and (b) in determining what material should receive protection and underscoring the social utility of effective tailoring of copyright law’s coverage).
other better-suited forms of legal protection.\textsuperscript{327} The criterion of the § 102(a) filter is broader and more open-ended: under the banner of “work of authorship,”\textsuperscript{328} § 102(a) requires a work to be of a general category or type of works for which copyright protection is appropriate.\textsuperscript{329} More detailed delimitation of the range of categorical inclusion is not left completely to courts, however. Instead, the section embodies an institutional division of labor.\textsuperscript{330} Section 102(a) lists subcategories with respect to which Congress explicitly pronounced judgments of copyright eligibility. But § 102(a) leaves courts to make judgments on whether other classes of works, such as sports moves, are appropriate for copyright, either by interpreting broadly the enumerated subcategories or by characterizing additional subject matter as falling within the residual category of “works of authorship.”\textsuperscript{331} In contrast, the originality requirement is a fine-grained, albeit weak filter. It calls for granular, case-by-case scrutiny of whether a particular work has enough of the quality of creativity deemed to be distinctively necessary for copyrightable subject matter—i.e. communicative, expressive forms.\textsuperscript{332}

The redundant doctrinal structure generated by these overlapping requirements helps circumvent the difficulty of formulating a single, reasonably administrable test for answering the question of what sorts of creations merit copyright protection. In this sense, the partially redundant copyrightability doctrines operate in a manner analogous to separate, two-dimensional maps that feature different focal points on a globe or different methodologies for projecting portions of the globe onto a two-dimensional surface. Together, these maps can provide a more readily usable and reliable method of envisioning the whole than any single, two-dimensional map—even one that comes with descriptions of its distortions and rules for stitching together its ends.

Finally, through the example of reproduction and derivative-work rights, we acknowledge that redundancy can sometimes go wrong, especially when decision-makers lose sight of the underlying purposes of relevant legal doctrines and their intended relationship. Even only partial redundancy can lead to sloppiness in distinguishing the proper character of overlapping doctrines, and at least occasionally, such sloppiness can produce

\textsuperscript{327} See supra text accompanying notes 232–34 (discussing the role of § 102(b) in filtering subject-matter better-suited for protection under patent law than copyright law).

\textsuperscript{328} 17 U.S.C. § 102 (2012).

\textsuperscript{329} See supra text accompanying notes 255–60 (explaining why copyright protection and proprietary rights in general may not be appropriate for certain types of subject matter).

\textsuperscript{330} See supra text accompanying notes 262–50 (explaining the “division of labor” between Congress and the courts).

\textsuperscript{331} See supra text accompanying note 265 (discussing the work that § 102(a) of the Copyright Act leaves to courts).

\textsuperscript{332} See supra text accompanying notes 251–54 (explaining the practical application of the originality requirement as well as its justifications).
confusion and a peculiarly perverse legal conclusion, such as we believe was the case in *RDR Books*. The point is independent of any view on the proper scope of copyright: even if one favors a broad scope that makes a wide range of derivative uses subject to copyright, derivative-work and reproduction rights should be understood in a way that is coherent in light of the policies supposedly served by each set of rights. To the extent redundancy leads to confusion, it can lead to muddled law and bad results.

But the example of *RDR Books* also illustrates how redundancy failure is not necessarily intrinsic to redundancy itself. Instead, a failure of redundancy to add net value rather than subtract it can reflect a failure by relevant decision-makers to pay attention to how redundant parts are meant to act and interact in concert. Of course, this means that deployment of partially redundant legal doctrines comes with information costs for actors who need to apply the law or predict its application. But this seems true whenever the law deploys a multiplicity of legal doctrines. Thus, information costs about the relationship and purposes of legal doctrines do not seem a special reason for rejecting redundancy wholesale, even if they might limit the number of legal doctrines that can be effectively deployed within any particular field.

In sum, copyright law and principles of redundancy and anti-redundancy each have something to teach the other. When analyzed in light of redundancy theory, various doctrinal features of copyright that might initially appear confused or wasteful, such as overlapping aspects of doctrines of copyrightability and fair use, turn out to embody apparently useful, functional design. Other overlapping doctrinal features, such as the provisions for reproduction and derivative-work rights, can lead to muddled analysis and questionable results consistent with pessimistic expectations of anti-redundancy theory.

Copyright law, for its part, enriches the study of redundancy in legal design. Copyright law exemplifies how redundancy in legal doctrine can advance substantial social ends by helping prevent undesirable case outcomes, by facilitating more expeditious dispute resolution, and by allowing individually administrable legal doctrines to together approximate a more complex whole. Distinct but overlapping doctrines of fair use, improper appropriation, originality, and copyrightable subject matter are likely to better police the boundaries of an expansive copyright regime than would a significantly more streamlined set of fully compartmentalized legal doctrines. In contrast, confusion with respect to the relationship between reproduction and derivative-work rights indicates how optimal operation of legal redundancy can entail information and administration costs, including those associated with continued attention to the underlying purposes of legal

---

333 See *supra* text accompanying notes 271–81 (explaining the court’s reasoning in *Warner Bros. Entm’t, Inc. v. RDR Books*, 575 F. Supp. 2d 513 (S.D.N.Y. 2008), and questioning the reasonableness of the outcome).
doctrines.

This Article has shown how redundancy can be a crucial strategy for protecting key social interests such as free speech interests. Like all strategies, however, redundancy is vulnerable to slipshod execution. With this caveat, legal policymakers might generally do well to incorporate redundancy as a basic, albeit not always controlling, principle for formulating and evaluating legal design.