ARTICLE

OF BASS NOTES AND BASE RATES: AVOIDING MISTAKEN INFERENCES ABOUT COPYING

Christopher Buccafusco & Rebecca Tushnet[*]

ABSTRACT

To prove copyright infringement, a plaintiff must convince a jury that the defendant copied from the plaintiff's work rather than independently creating it. To prove copying, especially in cases involving music, it's common for plaintiffs and their experts to argue that the similarities between the parties' creative works are so great that it is simply implausible that the defendant's work was created without copying from the plaintiff's work. Unfortunately, in its present form, the argument is mathematically illiterate; it assumes, without any underlying evidence, that the experts know or could reasonably estimate how likely it is that a song with similarity level x to another, earlier song was created without copying from the earlier song. Until the state of the underlying art changes, it is reasonable for experts to testify about the existence of similarities between works, but it is unsupported and unreasonable for them to testify about the likelihood that those similarities came about from copying. We don't know that likelihood in the absence of evidence about base rates: how common it is for a song to have similarity level x with some other song in the corpus of existing songs, and how common

it is for that similarity to come from copying or from independent creation (or from both songs copying a shared antecedent). Until that knowledge is available, testimony about the probability of copying should be deemed inadmissible under Federal Rule of Evidence 702.

TABLE OF CONTENTS

I. INTRODUCTION

Lana Del Rey's song "Get Free" sounds similar to Radiohead's song "Creep." But sounding similar is not equivalent to copyright infringement. The defendant must have actually copied her song from the plaintiff's song. So, did she? Or, did she copy The Hollies' 1974 tune "The Air I Breathe," to which her song also sounds similar?[1] In nonpiracy copyright infringement cases like this, especially those involving music, it's common for plaintiffs to argue that the similarities between the parties' creative works are so great that it is simply implausible that the defendant's work was created without copying from the plaintiff's work. Not all copying is unlawful, but this argument is often deployed both to prove the

---

1. As it turns out, The Hollies had previously sued Radiohead, and the writers of "The Air I Breathe" now own a substantial portion of the copyright in "Creep." Dee Lockett, *Making Sense of Radiohead's Nonsensical Copyright Dispute with Lana Del Rey*, VULTURE (Jan. 11, 2018), https://www.vulture.com/2018/01/radioheads-copyright-dispute-with-lana-del-rey-explained.html [https://perma.cc/B5VY-2HMJ].

fact of copying and to convince a factfinder that the copying was significant enough to constitute infringement.[2] And courts are sometimes sympathetic to these arguments.

The trouble is that, in its present form, the argument is mathematically faulty; it assumes, without any underlying evidence, that we can *know* how likely it is that a song with some degree of similarity to another, earlier song was created without copying from the earlier song. Our argument is modest: until the state of the underlying science changes, it is reasonable for experts to testify about the existence of similarities between works, but it is unsupported and unreasonable for them to testify about the likelihood that those similarities came about from copying. Experts don't know *that* likelihood in the absence of evidence about base rates: How common is it for a song to have similarity level *x* with some other song in the corpus of existing songs, and how common is it for that similarity to come from copying as opposed to from independent creation (or from both songs copying a shared antecedent)?[3]

It is conceivable that we could develop metrics to measure the first probability, and we could likely improve ways to think about the second one, but that work has not yet been done. Until it is, testimony about the likelihood that copying occurred based on shared similarities between two works fails to satisfy Federal Rule of Evidence (F.R.E.) 702 and should be excluded by judges.[4]

The problem is well-understood in other contexts. For example, where a serious disease is very rare, and there is a very accurate diagnostic test for it, one might think that everyone should be tested for it. But with universal screening for a rare disease with a low base rate (say 1 in 5 million actually have the disease), using a diagnostic test that is even slightly imperfect (say it generates "false positives" 1% of the time), it is mathematically

---

2.   *See* Shyamkrishna Balganesh et al., Essay, *Judging Similarity*, 100 IOWA L. REV. 267, 284–86 (2014).

3.   For a longer discussion of copying and independent creation, see Christopher Buccafusco, *There's No Such Thing as Independent Creation, and It's a Good Thing, Too*, 64 WM. & MARY L. REV. 1617, 1628–33 (2023); see also JENNIFER JENKINS, MUSIC COPYRIGHT, CREATIVITY, AND CULTURE (forthcoming Feb. 2024) (manuscript at 104–15) (draft on file with authors); Christopher Brett Jaeger, Note, *"Does That Sound Familiar?": Creators' Liability for Unconscious Copyright Infringement*, 61 VAND. L. REV. 1903, 1912–14 (2008).

4.   *See* FED. R. EVID. 702 (setting forth the criteria by which an expert may qualify to testify to their opinion). Alfred Yen is independently developing a similar line of arguments in a new working paper. *See* Alfred C. Yen, The Evidentiary Use and Misuse of Forensic Musicology in Copyright Litigation 1, 3–4 (July 27, 2023) (unpublished manuscript) (on file with authors).

certain that most positive results from the test will be false positives. Thus, a person who tests positive for the disease is more likely than not to *not* have the disease, even though the test is almost perfectly accurate. It is for this reason that public health professionals generally don't recommend universal screening for very rare diseases.[5] Many people who hear that a 99% accurate test was positive for a fatal fetal anomaly understandably believe that the result means that it is overwhelmingly likely that the fetus is afflicted.[6] But, unless screening is limited to those already known through other means to be at risk, it is often quite *unlikely* that the fetus is afflicted.

Our argument is not that we know the base rate of similarity or the accuracy of experts' assessments of how often similarity results from copying (false positives). To the contrary, it is that we *don't* know. Similarity makes copying more likely than its absence does, but similarity does not itself make copying more likely than not. Until experts can say more about base rates and rates of copying, they should be saying less. It follows that courts should exclude testimony that relies on unsupported assumptions about probabilities.

Part I provides a short introduction to copyright law's doctrine related to proving copying and provides several examples of expert testimony in music copyright cases involving the probability of copying. Part II explains base-rate neglect, and then shows how the same problems arise in the context of copyright litigation. Part III considers various routes forward.

II. PROVING COPYING

A plaintiff in a copyright infringement case must do more than prove that the defendant's work is similar to its work. The plaintiff must also prove that the defendant copied from the

---

5. *See* Amanda Fakih & Kayte Spector-Bagdady, *Should Clinicians Leave "Expanded" Carrier Screening Decisions to Patients?*, 21 AMA J. ETHICS 858, 859–61 (2019).

6. Sarah Kliff & Aatish Bhatia, *When They Warn of Rare Disorders, These Prenatal Tests Are Usually Wrong*, N.Y. TIMES (Jan. 1, 2022), https://www.nytimes.com/2022 /01/01/upshot/pregnancy-birth-genetic-testing.html [https://perma.cc/85LV-K8BZ]; *Genetic Non-Invasive Prenatal Screening Tests May Have False Results: FDA Safety Communication*, U.S. FDA (Apr. 19, 2022), https://www.fda.gov/medical-devices/safety-com munications/genetic-non-invasive-prenatal-screening-tests-may-have-false-results-fda-saf ety-communication [https://perma.cc/Z5AA-LD2S]; *In re* Natera Prenatal Testing Litig., No. 22-cv-00985-JST, 2023 WL 3370737, at *1–2 (N.D. Cal. Mar. 28, 2023) (explaining consumer protection claims based on alleged misleadingness of touting accuracy of tests when low base rates make false positive results more likely than true positive results).

plaintiff rather than independently creating its work.[7] This doctrinal requirement is known as "copying-in-fact."[8] Sometimes, proving copying is easy because the defendant readily admits to it.[9] But, in many cases, proving copying can be challenging and plaintiffs must rely on circumstantial evidence to establish it. In this Part, we explain the doctrine on proving copying-in-fact, and we describe how plaintiffs often resort to providing expert evidence about the probability of similarities arising from copying rather than independent creation.

A. *Copying-in-Fact Doctrine*

Copying-in-fact is an empirical question: did the defendant copy from the plaintiff or, rather, did the defendant create independently of the plaintiff?[10] In some cases, proving the empirical reality is simple. For example, Robin Thicke and Pharrell Williams admitted that they had Marvin Gaye's music in mind when they wrote "Blurred Lines."[11] However, defendants often claim that they created their works independently, never having experienced the plaintiff's work before.[12] In such cases, plaintiffs must rely on circumstantial evidence of copying.[13]

---

7. Skidmore v. Led Zeppelin, 952 F.3d 1051, 1064 (9th Cir. 2020) (en banc) (first citing Rentmeester v. Nike, Inc., 883 F.3d 1111, 1116–17 (9th Cir. 2018); then citing Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345–46 (1991)).

8. Buccafusco, *supra* note 3, at 1629–30.

9. For examples of defendants admitting to instructing their designers to replicate the plaintiffs' work, see Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp., 562 F.2d 1157, 1165 (9th Cir. 1977) and Steinberg v. Columbia Pictures Indus., Inc., 663 F. Supp. 706, 710 (S.D.N.Y. 1987).

10. Here, "independently" means only that the defendant didn't copy from the plaintiff; perhaps the defendant copied from something else, including something from which the plaintiff also copied. *See, e.g.*, Adam Neely, *Did Dua Lipa ACTUALLY Plagiarize Levitating?*, YOUTUBE (Mar. 6, 2022), https://www.youtube.com/watch?v=HnA1QmZvSNs [https://perma.cc/LEC4-6QYN] (suggesting that both plaintiff and defendant in *Cope v. Warner Records, Inc.* might have been inspired by another artist's hit).

11. *See* Stelios Phili, *Robin Thicke on that Banned Video, Collaborating with 2 Chainz and Kendrick Lamar, and His New Film*, GQ (May 6, 2013), https://www.gq.com/story/robin -thicke-interview-blurred-lines-music-video-collaborating-with-2-chainz-and-kendrick-lam ar-mercy [https://perma.cc/9WEZ-9G5T] (stating Thicke suggested to Williams that they should create music like Gaye's "Got to Give It Up").

12. *See* Sheldon v. Metro-Goldwyn Pictures Corp., 81 F.2d 49, 52, 54 (2d Cir. 1936) ("[I]f by some magic a man who had never known it were to compose anew Keats's Ode on a Grecian Urn, he would be an 'author,' and, if he copyrighted it, others might not copy that poem, though they might of course copy Keats's.").

13. Rentmeester v. Nike, Inc., 883 F.3d 1111, 1117 (9th Cir. 2018).

Recently, federal courts have clarified the elements of a plaintiff's case for proving copying-in-fact.[14] First, the plaintiff must prove that the defendant had *access* to the plaintiff's work, i.e., that the defendant probably experienced the plaintiff's work in the past.[15] Though some courts excuse the requirement of showing access with evidence other than the two works themselves if their similarities are so "strikingly similar as to preclude the possibility of independent creation."[16] Second, the plaintiff must prove that the defendant's work exhibits *similarities probative of copying*.[17]

The intuition behind the access prong is clear: if the defendant did not have access to the plaintiff's work, she couldn't have copied it. Lack of access implies independent creation, while proof of access creates the possibility of copying.[18] Plausibly alleging access is often straightforward, especially in the Internet era.[19] For example, when a group of Christian hip-hop artists sued Katy Perry for copyright infringement, the court didn't find it necessary to address access even though the plaintiffs' song was not successful outside of the Internet and was not played on the radio.[20] It was sufficient that the song had received millions of views on YouTube, presumably one of which could have been Katy Perry or her co-authors.[21]

---

14.   *See* Skidmore v. Led Zeppelin, 952 F.3d 1051, 1067–69 (9th Cir. 2020) (en banc) (overruling the inverse ratio rule because it unfairly favors popular, and therefore, more accessible, works); *Rentmeester*, 883 F.3d at 1117 (explaining that in the absence of direct evidence of copying, the plaintiff can show access and similarities probative of copying).

15.   *Rentmeester*, 883 F.3d at 1117–18.

16.   Lipton v. Nature Co., 71 F.3d 464, 471 (2d Cir. 1995) (quoting Ferguson v. NBC, Inc., 584 F.2d 111, 113 (5th Cir. 1978)). In one recent case, for example, the court found that the plaintiffs hadn't plausibly pled access but could still proceed to discovery about Dua Lipa's hit "Levitating," based on an allegedly striking similarity of "repetitive rhythm" and "signature melody." Larball Publ'g Co. v. Lipa, No. 22 Civ. 1872 (KPF), 2023 WL 5050951, at *13–15 (S.D.N.Y. Aug. 8, 2023). As the text will explain *infra*, there is simply no basis for inferring copying based on fragmented musical similarities of this type. Indeed, as a music theorist explained with respect to another lawsuit against "Levitating" for similar similarities, the challenged aspects use the well-known Charleston rhythm. Neely, *supra* note 10.

17.   *Led Zeppelin*, 952 F.3d at 1064 (quoting Rentmeester v. Nike, Inc., 883 F.3d 1111, 1117 (9th Cir. 2018)).

18.   Granite Music Corp. v. United Artists Corp., 532 F.2d 718, 721 (9th Cir. 1976).

19.   *See* Design Basics, LLC v. Lexington Homes, Inc., 858 F.3d 1093, 1106–08 (7th Cir. 2017) (listing a handful of district court cases that have found access based on the plaintiff's internet presence); *see also* Carissa L. Alden, Note, *A Proposal to Replace the Subconscious Copying Doctrine*, 29 CARDOZO L. REV. 1729, 1731–32 (2008).

20.   Gray v. Hudson, 28 F.4th 87, 92–93, 96 (9th Cir. 2022).

21.   *Id.* at 93.

Merely having access to a work doesn't prove that the defendant copied it. Katy Perry obviously had access to Beethoven's 9th Symphony, but no one thinks that "Dark Horse" copied it. That's because the two works are not similar. Thus, a plaintiff must prove not just access to her work but also that the defendant's work exhibits similarities to the plaintiff's work that would lead us to believe that it was copied from the plaintiff. As the en banc Ninth Circuit explained in its recent *Led Zeppelin* holding: "'similarities probative of copying' . . . show that the similarities between the two works are due to 'copying rather than . . . coincidence, independent creation, or prior common source.'"[22] Or, as Judge Richard Posner explained, the question is whether the similarities between the two works make the factfinder "suspicious" of copying.[23]

The probative similarity element presents an empirical challenge. The question is: are the similarities between the plaintiff's and defendant's work the kinds of similarities that probably arise due to copying or, instead, are they the kinds of similarities that might arise even though the defendant independently created its work? Consider, for example, the mannequins illustrated in Figure 1:

---

22.   *Led Zeppelin*, 952 F.3d at 1064 (first quoting *Rentmeester*, 883 F.3d at 1117; then quoting Bernal v. Paradigm Talent & Literary Agency, 788 F. Supp. 2d 1043, 1052 (C.D. Cal. 2010)).

23.   *See* Ty, Inc. v. GMA Accessories, Inc., 132 F.3d 1167, 1170 (7th Cir. 1997).

Figure 1[24]

Mara   Liza

Obviously, both mannequins will share certain features merely because they both represent female faces—the shape of their eyes, noses, and mouths might be similar simply because this is what some people look like. These similarities exist because both mannequins are copying features of real (or imagined) public domain humans. But other similarities might be suggestive of copying. In particular, the plaintiff's mannequin accidentally had two hairlines, because the initial hairline was placed in the wrong location.[25] When we learn that the defendant's mannequin also has two hairlines—a feature not shared by real humans nor one that is otherwise part of the public domain—we are compelled to infer that the defendant copied from the plaintiff. There is no other reason for both mannequins to share this feature.

At the same time, coincidences do happen, such as two photographers simultaneously capturing images that look the

---

24.   Pivot Point Int'l, Inc. v. Charlene Prods., Inc., 170 F. Supp. 2d 828, 840 (N.D. Ill. 2001), *rev'd*, 372 F.3d 913 (7th Cir. 2004).

25.   Pivot Point Int'l, Inc. v. Charlene Prods., Inc., 372 F.3d 913, 915–16 (7th Cir. 2004).

same (this has happened more than once),[26] two writers of fantasy for children naming a glasses-wearing protagonist H/Larry Potter and confronting him with Muggles,[27] and two science fiction movies coming up with the name Ewoks for short, furry, arboreal aliens.[28] Given that humans share many cultural and physiological features, so that what "sounds right" or "looks right" will often be shared across creators, and given the generation of millions of works of creativity, at least some instances of coincidental striking similarity are inevitable.[29]

Ultimately, circumstantial proof of copying-in-fact is, like all circumstantial evidence, about the strength of the inferences that can be drawn from the facts. Some works are similar because they both draw from sources in the public domain, whether specific (a moment in time) or general ideas (a generic name for a white, English-speaking boy, a signifier of nerdiness, and a word that sounds unusual but unthreatening). Other works are similar because the defendant copied from the plaintiff. The question for the court, then, is whether the similarities at issue fall into the former or the latter category. Plaintiffs routinely allege that similarity is so unusual and extreme that copying is the only plausible explanation.[30]

---

26. Ron Risman, *How Two Photographers Unknowingly Shot the Same Millisecond in Time*, PETAPIXEL (Mar. 7, 2018), https://petapixel.com/2018/03/07/two-photographers-un knowingly-shot-millisecond-time/ [https://perma.cc/Y5RW-E94L]; Michael Zhang, *Contest Copyright Controversy a Crazy Coincidence*, PETAPIXEL (Feb. 3, 2015), https://petapixel.com/20 15/02/03/contest-copyright-controversy-crazy-coincidence/ [https://perma.cc/R35B-LMQT].

27. Scholastic, Inc. v. Stouffer, 221 F. Supp. 2d 425, 427–30 (S.D.N.Y. 2002).

28. Preston v. 20th Century Fox Can. Ltd., 1990 CarswellNat 205, para. 7–9 (Can.) (WL). As described by Cameron Hutchison, "striking similarities" between the Ewoks in plaintiff's *Space Pets* and the Ewoks in *Return of the Jedi* include not just the name but a portrayal of a small, bipedal, forest-dwelling species that "lives in tree villages with connecting bridges/swinging vines; [uses] net vine traps and sedan chairs . . . ; speak[s] in high squeaky voices; and [uses] a language machine translator." Cameron Hutchison, *Insights from Psychology for Copyright's Originality Doctrine*, 52 IDEA 101, 105–06 (2012). Nonetheless, the defendants persuaded the court that these similarities resulted not from copying but from creators drawing from the same cultural well. *Id.* at 112–13.

29. *Cf.* Jake Linford, *Are Trademarks Ever Fanciful?*, 105 GEO. L.J. 731, 749–51 (2017) (using psychological and linguistic literature to argue that many "meaningless" sounds, images, etc., are inherently attractive for expressing particular ideas in a given cultural setting).

30. *See, e.g.*, Complaint for Copyright Infringement & Demand for Jury Trial at 4, Cope v. Warner Records, Inc., No. 2:22-cv-01384 (C.D. Cal. Mar. 1, 2022) ("Given the degree of similarity, it is highly unlikely that 'Levitating' was created independently from 'Live Your Life.'").

B.   *Expert Evidence of Copying*

Circumstantial evidence of copying presents challenging empirical questions for courts and jurors. Unsurprisingly, then, plaintiffs in music cases often rely on expert witnesses to convince the jury that the defendant copied from the plaintiff. Consider the following examples. In each of them, the expert musicologist offers evidence on the likelihood that the defendant copied from the plaintiff.

In *Selle v. Gibb*, the Bee Gees were accused of copying a song written by a department store clerk that had never been published or performed on the radio.[31] Dr. Parsons, the plaintiff's expert, testified that, in his opinion, the two songs had such striking similarities that they could not have been written independently of one another.[32] He also testified that he did not know of two songs by different composers "that contain as many striking similarities" as did the two songs at issue there.[33] The Bee Gees didn't provide any expert evidence to contradict Dr. Parsons, and the jury held them liable for copyright infringement.[34]

In *Gaste v. Kaiserman*, the plaintiff, a French citizen, alleged that the 1973 hit song "Feelings" by Morris Albert was copied from his 1956 song "Pour Toi," which was released as part of an unpopular French movie that never achieved wide distribution.[35] Gaste's expert told the jury that "there is not one measure of 'Feelings' which . . . cannot be traced back to something which occurs in 'Pour Toi.'"[36] His evidence also included "a unique musical 'fingerprint'—an 'evaded resolution'—that occurred in the same place in the two songs."[37] Gaste's expert explained, "that while modulation from a minor key to its relative major was very common, he had never seen this particular method of modulation in any other compositions."[38] According to the plaintiff's expert, it would be impossible to compose "Feelings" without copying from "Pour Toi."[39] This time, the defendants presented their own expert musicologist to challenge the plaintiff's expert, but the jury still

---

31.   Selle v. Gibb, 567 F. Supp. 1173, 1175–76 (N.D. Ill. 1983).

32.   *Id.* at 1177–78.

33.   *Id.* at 1178.

34.   *Id.* at 1179.

35.   Gaste v. Kaiserman, 863 F.2d 1061, 1063 (2d Cir. 1988).

36.   *Id.* at 1068.

37.   *Id.*

38.   *Id.*

39.   *Id.* at 1069.

found infringement and the Second Circuit upheld their verdict on appeal.[40]

To take a more recent example, consider the lawsuit against Ed Sheeran's "Thinking Out Loud" by the owners of Marvin Gaye's interest in the copyright to "Let's Get It On"—this is only one of several recent infringement lawsuits against the song.[41] The complaint includes two separate extended analyses by the plaintiff's musicologists. Professor John Covach notes several "melodic similarities" between the two songs. He argues that:

> It is possible that, taken in isolation, one might view these melodic similarities as a mere coincidence, arising from the limitation of the musical materials available within the pop style. But in light of the other marked musical parallels discussed here, the sense that these melodic passages in "Thinking Out Loud" create an additional set of references to "Let's Get It On" is vastly increased. In the fullest musical and music-analytical context, it seems almost impossible— or at least impossibly unlikely—that these similarities arose by simply by chance.[42]

For his part, Professor Walter Everett begins by noting:

> In judging the dependencies between two or more songs, one must consider the combinations of factors at play and the degrees of their similarity or difference, measured against the commonness or rarity of these factors in the general population (particularly in the population of typical sources, being pop music produced before 1980).[43]

Everett conducted an analysis of over 6,000 songs between 1955 and 1975 and found only a handful that included the features he considers most distinctive in "Let's Get It On."[44] From this he asserts:

> It has been shown above that the combination of two or three of the above factors occurs *only* in these two songs. The fact that all factors are held in common proves the case . . . . The additional fact that no other song shares the foregoing characteristics is strong evidence not of arbitrary

---

40.	*Id.* at 1068, 1071.

41.	Structured Asset Sales, LLC v. Sheeran, No. 18 Civ. 5839, 2023 WL 3475524, at *1 (S.D.N.Y. May 16, 2023) (explaining that this case is one of several against Sheeran for copying the Gaye song, and it is different from the one that a jury recently found to be non-infringing).

42.	Complaint at 37, Structured Asset Sales, LLC v. Sheeran, No. 1:20-cv-4329 (S.D.N.Y. June 8, 2020).

43.	*Id.* at 40–41.

44.	*Id.* at 57–59.

convergence or parallel development, but knowing, willful infringement.[45]

In each of these cases, plaintiffs' experts made empirical assertions about the strength of the inference of copying based on similarities between the works and the rarity or distinctiveness of those similarities. Notice, however, that none of these assertions contain any meaningful quantification. Professor Everett in the Ed Sheeran litigation comes the closest, referring to a sample of 6,000 songs, only a few of which share the same features. Yet as we explain in the next Part, the inference of copying is deeply dependent on establishing plausible, quantifiable empirical foundations. The above examples are virtually worthless in their ability to establish the probability of copying.

III. BASE-RATE NEGLECT IN MUSIC COPYRIGHT

A.   *Understanding Base-Rate Probabilities*

The basic problem with this sort of circumstantial evidence is well understood by statisticians, though hardly by anyone else.[46] Consider, for example, a proposal for mandatory Human Immunodeficiency Virus (HIV) testing for the entire U.S. population. HIV tests are very accurate on an individual level. That is, they are very likely to correctly identify someone with HIV (a measure known as sensitivity: the test will correctly detect 98.3% of people who carry HIV, so the percentage of false negatives is low) and very unlikely to incorrectly identify someone without HIV (a measure known as specificity: the test will correctly detect 99.8% of people who do not carry HIV, so the percentage of false positives is low).[47] Intuitively, testing everyone thus makes sense because the resulting information should be very reliable.

But one number is missing from this intuition: the base rate or underlying prevalence of HIV in the population as a whole. The chance that a positive test actually signifies HIV infection—known as the positive predictive value—depends heavily on how many

---

45.   *Id.* at 66.

46.   *See* KENNETH R. FOSTER & PETER W. HUBER, JUDGING SCIENCE: SCIENTIFIC KNOWLEDGE AND THE FEDERAL COURTS 121 (1999); Brandon Garrett & Gregory Mitchell, *How Jurors Evaluate Fingerprint Evidence: The Relative Importance of Match Language, Method Information, and Error Acknowledgment*, 10 J. EMPIRICAL L. STUD. 484, 496 (2013).

47.   Paul D. Cleary et al., *Compulsory Premarital Screening for Human Immunodeficiency Virus: Technical and Public Health Considerations*, 258 JAMA 1757, 1758 (1987).

people actually have HIV.[48] If 1 in 3,000 people have the virus, then those very small percentages of false positives will add up to big numbers—bigger than the large percentages, but small absolute numbers, of true positives. With such a low base rate, a randomly selected person who tests positive for HIV has less than a 15% chance of actually being infected.[49] True, that's likelier than 1 in 3,000, but it's much less likely than a false positive, and may well not be helpful enough for its probative value to outweigh its prejudicial effects on factfinders who predictably struggle to understand how such a reliable test could produce unreliable results.[50]

This failure to consider how common it is for the underlying fact of interest to occur is known in psychological literature as the "base-rate fallacy"[51] or "base-rate neglect."[52] According to the behavioral economists Daniel Kahneman and Amos Tversky, "[t]he failure to appreciate the relevance of prior probability in the presence of specific evidence is perhaps one of the most significant departures of intuition from the normative theory of prediction."[53]

Sometimes, as in the HIV example, we can know the underlying base rate.[54] However, in other situations, we have to estimate the base. Importantly, small changes in our assumptions about base rates can cause big changes in positive predictive value. Consider an example drawn from the famous torts case of *Byrne v. Boadle*.[55] A warehouse maintains barrels of flour, one of which broke loose, rolled out of a window, and fell on the plaintiff's head.[56] If a barrel of flour is improperly handled, it will break loose

---

48.   *See id.* at 1758.

49.   *Id.* at 1759–60; *see* Neal C. Stout & Peter A. Valberg, *Bayes' Law, Sequential Uncertainties, and Evidence of Causation in Toxic Tort Cases*, 38 U. MICH. J.L. REFORM 781, 789 (2005) (offering further examples and demonstrating that prevalence substantially affects whether a positive result is more likely correct or incorrect).

50.   *See* Boaz Sangero, *Safety from Flawed Forensic Sciences Evidence*, 34 GA. ST. U. L. REV. 1129, 1145–46 (2018) (explaining further that a person from a low-risk group will be even more likely to get a false positive result).

51.   Maya Bar-Hillel, *The Base-Rate Fallacy in Probability Judgments*, 44 ACTA PSYCHOLOGICA 211, 212 (1980); *see also* Amos Tversky & Daniel Kahneman, *Evidential Impact of Base Rates*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 153–54 (Daniel Kahneman et al., eds., 1982).

52.   Tversky & Kahneman, *supra* note 51.

53.   Daniel Kahneman & Amos Tversky, *On the Psychology of Prediction*, 80 PSYCHOL. REV. 237, 243 (1973).

54.   *See supra* text accompanying notes 46–48.

55.   Byrne v. Boadle (1863) 159 Eng. Rep. 299, 299; 2 H & C 722, 722.

56.   Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 808 (2001) (citing *Byrne*, 159 Eng. Rep. 299).

90% of the time. Yet, if a barrel is properly handled, it only breaks loose 1% of the time. In this warehouse, barrels are properly handled 99.9% of the time: the base rate of improper handling is quite low.[57] How likely is it that the barrel of flour that hit the plaintiff was improperly handled?

The answer is about 8%.[58] But that's not what most judges think, according to a study by Guthrie, Rachlinski, and Wistrich.[59] Many of the judges in their study thought that the probability of negligence was at least 25%, and a full 40% of them thought the answer fell in the range of 76%–100%.[60] While a reasonable percentage of judges in the study did reach the correct answer, these results are troubling.[61]

Note, though, how quickly the probability of negligence changes with even a small change in the base rate. If instead the warehouse properly handles the barrels only 99.0% of the time rather than 99.9% of the time, the answer changes dramatically. Now the plaintiff's injury is about 47% likely to have been caused by improper handling.[62]

Why should this matter when the plaintiff's expert is testifying about a particular instance of similarity? Because the expert is testifying about this instance based on her expertise in other instances—she is saying that, in her experience, this kind of similarity must result from copying. That is, every expert statement about the unlikelihood of independent creation based on the extent of the similarity between two works is based on the underlying thesis that we should use generalizations about creation in the genre to evaluate this individual example.[63]

---

57. *Id.* at 809.

58. *Id.* Assume 100,000 handlings:

99,900 proper x 0.01 = 999 barrels breaking loose

100 improper x 0.90 = 90 barrels breaking loose

Thus, about 8% are due to improper handling (90/1089 = 0.083)

59. *Id.*

60. *Id.*

61. About 40% of the judges selected the option for 0–25% likely. *Id.*

62. Again, assume 100,000 handlings:

99,000 proper x 0.01 = 990 barrels breaking loose

1,000 improper x 0.90 = 900 barrels breaking loose

Thus, about 47% are due to improper handling (900/1890 = 0.476)

63. *See* FOSTER & HUBER, *supra* note 46 ("Statisticians understand this 'base-rate' problem very well. After *Daubert* [509 U.S. 579 (2003)], judges must grasp it too. This is absolutely fundamental to the evaluation of the reliability of a claim based on an observation or a test of any kind.").

Outside of copyright, courts have sometimes recognized the problem of base-rate neglect—indicating that they could do better in copyright as well.[64] In *Gonzalez v. Metropolitan Transportation Authority*, the Ninth Circuit Court of Appeals explained that random drug testing generated the same problem:

> [Suppose] an error rate such that one person out of 500 gets a report of "dirty" urine when it was actually "clean." Suppose that . . . on any particular day one worker in 10 has alcohol or drugs in his blood. Then with a 1/500 false positive rate, out of 1,000 tests, 2 will be positive even though the employee's urine was clean, and 100 will be positive correctly. Only one of the positives out of every 51 is false. . . . That is a fairly effective test, in terms of reliability.
>
> But if the workers are generally "clean," the reliability of the test goes way down. Suppose on a particular day only one worker in 500 has ingested drugs or alcohol. Then with a 1/500 false positive rate, out of 1,000 tests, 2 will be correct positives and 2 will be false positives. Half the employees who get a "dirty" urinalysis report are unjustly categorized. A positive result is as likely to be false as true on so clean a population, even though the test is identical to the one that was quite effective for a population with a higher incidence of drug and alcohol usage.[65]

---

64.   Not always. In *Lee v. Martinez*, 96 P.3d 291, 302–03 (N.M. 2004), the New Mexico Supreme Court rejected the argument that an unknown base rate made polygraph results unreliable. Unfortunately, the court did so by making assumptions about the base rate, reasoning that, whether the base rate of lying was 50% or 90%, the polygraph results would still provide relevant evidence. *Id.* The court stated that:

> [T]he base rate has no effect on the reliability of the polygraph—regardless of whether 50% or 90% of the sample population is deceptive, the accuracy of the polygraph remains unchanged. The base rate only affects the confidence that we have in making decisions based on the results of any one polygraph examination.

Id. at 302.

Notably, the court didn't consider a full range of base rates—at a base rate of 99% lying, the polygraph is much more likely to wrongly exonerate a liar than to identify a truth-teller, and at a base rate of 99% truthful, the polygraph is much more likely to falsely implicate a truth-teller than to identify a real liar. *See id.* The court reasoned that, as to any given individual, the fact that they passed the polygraph would allow us to update our beliefs beyond the base rate: with a 90% base rate of lying and a test with 90% sensitivity and specificity, "[p]rior to the subject passing the polygraph examination, we would have assumed only a 10% chance that subject was truthful. After passing the examination, though, the likelihood the subject was truthful has increased to 50%." *Id.* But note the sleight of hand: the court added back in the assumption that the base rate was known, even though it concededly was not; without that assumption, there is no way to tell how much to update our beliefs, and thus, quantitative claims are inappropriate. *See id.*

65.   Gonzalez v. Metro. Transp. Auth., 174 F.3d 1016, 1023 (9th Cir. 1999).

The issue has also occurred in other intellectual property disputes. In *Pharmacia Corp. v. Alcon Laboratories, Inc.*, Pharmacia made a glaucoma drug, Xalatan, and the defendant, Alcon, introduced a competing drug, Travatan.[66] The plaintiff sued for trademark infringement and submitted a statistical model based on reported instances in which one drug had been confused with another in medical practice.[67] Using this model, Pharmacia argued that the two drugs in the suit were likely to be confused with each other given their lexical similarities to those "error pairs."[68] This sounded convincing—convincing enough that the court denied a motion to exclude the expert's report—but, on examination, it revealed the same base-rate problem.[69]

Pharmacia's expert, Dr. Lambert, developed his model by using 1,127 "[k]nown [e]rror [p]airs" and then generating a list of 1,127 "[n]on–[e]rror [p]airs" by replicating the drug names in the first set.[70] Measures of lexical similarity were good at predicting which bucket a given pair from that group of 2,254 pairs fell into. But the expert himself identified over 235 million name pairs as the actual universe of potential brand name drug pairs—five orders of magnitude greater.[71] The base rate (the likelihood that a given pair of names was an error pair) was unknown but, based on simple calculations of the number of drug names in existence and the number of medication errors, extremely low. As a result, although the model had a sensitivity of over 93%, over 99 of every 100 predicted error pairs would be false positives.[72]

The relevant point here is that the expert's claims were unsupported and, in the present state of knowledge, unsupportable. We simply did not know how many drug error pairs were out there but unreported, just as we simply don't know the prevalence of striking similarities between two otherwise different works, much less how often those similarities are coincidental or result from copying. As the *Pharmacia* court concluded, the "inability to settle on a particular number (or even

---

66. Pharmacia Corp. v. Alcon Lab'ys, Inc., 201 F. Supp. 2d 335, 339–40 (D.N.J. 2002).

67. *Id.* at 339–40, 356–57.

68. *Id.* at 357.

69. *See id.* at 359 (explaining that the expert's model likely grouped the trademarks in question together with the group of false positives, instead of linking it to the correct "error pairs").

70. *Id.* at 358 & n.3.

71. *Id.* at 358.

72. *Id.* at 358–59.

range) not only creates great uncertainty, but undermines the statistical foundation and reliability of [the] model."[73]

We very much need to know the base rate of the behavior of interest (whether drug use, confusion between two medications based on their names, or copying) before we know if even very sensitive standards or tests are likely to be right in a given instance. But it is mathematically certain that the lower the base rate, the less likely such tests are to be correct.[74]

Given the clarity of the math, why are claims of striking similarity so intuitively persuasive as "proof" of copying? One way to think about it is that people often confuse two questions. The first is an actual legal question in copyright infringement cases: Given the evidence (similarities), how likely is it that the relevant event actually happened (copying)? The second question is: If we assume there was copying, how likely is it that these similarities would be present? Or, analogously, (1) given a positive test result, how likely is it that a person has the tested-for condition, in contrast to (2) given that a person has the condition, how likely is it that they would test positive? The second question is often

---

73.   *Id.* at 359. The court also pointed out that, given the unknowns, the predictive value of the model was dependent on arbitrary guesses. *Id.* at 360. For example, setting the number of known error pairs to 10,000 and assuming that there were only 2 million pairs of drug names, the model would still be wrong more than 90% of the time—a dramatic increase, but not good enough for reliability. *Id.* Assuming that the prevalence (base rate) of error pairs was up to 10%, the positive predictive value of the model improved to 62%, but there was no evidence that those were the right numbers, and the court concluded that, "because his model's predictive value turn[ed] entirely on initial assumptions about how many pairs should be included in the model, . . . [the] entire methodology [was] suspect." *Id.* This is a striking litigated example of the ways in which claims about likelihood can be misleading if factfinders ignore base rates; after being educated on the math, the court understood that "basic statistical principles" proved that the initially convincing claim of 99% likelihood of confusion was almost certainly a false positive. *Id.* at 378. It simply did not support the claim for which it was offered as proof. *Id.* (citing Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993)) (stating that courts should consider the error rate of scientific techniques); Kumho Tire Co. v. Carmichael, 526 U.S. 137, 153–54 (1999) (noting that the question is not whether a method is useful in general, but whether a method is reasonable in reaching a conclusion about the specific event at issue); Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997) ("A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered.").

74.   *See* REBECCA G. KNAPP & M. CLINTON MILLER III, CLINICAL EPIDEMIOLOGY AND BIOSTATISTICS 37–38, 40 (1992) (explaining the direct relationship between predictive value and base rate: lower prevalence means lower predictive value, and so too with higher prevalence); Stout & Valberg, *supra* note 49, at 786 ("Other things being equal, the lower the rate of the agent-induced disease among those with the disease, the less reliable the proffered causation opinion will be. Stated positively, the lower the rate of the agent-induced disease among those with the disease, the more sensitive and specific the tests employed by the expert need to be to achieve the level of reliability required for legal proof of causation.").

cognitively easier and people may therefore answer it instead, never noticing that they have fundamentally reversed the question.[75] This is why the logical error is sometimes known as the inverse fallacy; people have a tendency to treat the probability of a hypothesis given the evidence as the same as, or close to, the probability of the evidence given the hypothesis.[76]

Base-rate neglect is also related to availability bias: the tendency of humans to treat narrative coherence as evidence of causation.[77] The plaintiff tells a story in which copying occurred; if that story is narratively satisfying, it may prove persuasive despite the fact that the similarities were coincidental—after all, isn't it *suspicious* to have Larry Potter and the Muggles in one fantasy book and Harry Potter and the Muggles in another? But one thing that many people neglect is that the litigation context is not random; of the hundreds of thousands of songs (and books) published each year, only a few are successful, and there are millions of potential works to then be compared to those successful songs. With these odds, coincidence will eventually produce some surprising similarities. This is one reason that the mannequin case is plausibly different: the rate of mannequin production is nowhere as high.[78] Potential plaintiffs can pick and choose, making coincidence into conspiracy, aided by the fact that most people are not good at assessing probabilities.[79]

---

75. *See* Tversky & Kahneman, *supra* note 51, at 84–85, 98.

76. Guthrie et al., *supra* note 56, at 807.

77. *See* Hadar Y. Jabotinsky, *Revolving Doors–We Got It Backwards*, 89 U. CIN. L. REV. 432, 445–46 (2021) (stating the availability asserts that individuals calculate the probability of an event based on how easily individuals can recall similar past events).

78. See *Mannequin Market Size, Share & COVID-19 Impact Analysis*, FORTUNE BUS. INSIGHTS, https://www.fortunebusinessinsights.com/mannequin-market-106818 [https://pe rma.cc/FEX8-PVS4] (last visited Aug. 16, 2023) (stating that the global mannequin market was estimated to be worth over five billion dollars in 2022; not breaking out how many had faces or how many different faces there were); *Music Market Size & Share Analysis - Growth Trends & Forecasts (2023–2028)*, GLOBENEWSWIRE (Aug. 15, 2023, 6:02 PM), https://ww w.globenewswire.com/news-release/2023/08/15/2725967/0/en/Music-Market-Size-Share-An alysis-Growth-Trends-Forecasts-2023-2028.html [https://perma.cc/QB8V-RJ86] (stating that the size of the music industry in 2023 was estimated to be over twenty-eight billion dollars).

79. *See* Michael Shermer, *Why Our Brains Do Not Intuitively Grasp Probabilities*, SCI. AM. (Sept. 1, 2008), https://www.scientificamerican.com/article/why-our-brains-do-not-intuitively-grasp-probabilities/ [https://perma.cc/2M9Z-BNQD]. This problem is related to the birthday paradox: although it sounds highly unlikely for two randomly chosen people to have the same birthday, in a group of merely twenty-three, the odds are roughly even that there will be a match—because each person is actually being compared with everyone else in the group, not just one randomly selected person. Science Buddies, *Probability and the Birthday Paradox*, SCI. AM. (Mar. 29, 2012), https://www.scientificamerican.com/artic le/bring-science-home-probability-birthday-paradox/ [https://perma.cc/7YD6-CLED]. The

The core point is this: you can't say how likely it is that *B* (similarity) is the result of *A* (copying) without knowing the prior probability of *B* given *A* (similarity given copying) and of *B* given *not-A* (similarity given no copying), and the distribution of *A* and *not-A* in the population. Even if you assume that copying always results in similarity, the other numbers remain to be supplied, or at least estimated in some minimally reliable way.

B. *Assessing Copying in Music Copyright Litigation*

Let's consider what it would mean to estimate these numbers in music copyright cases. Assume that the plaintiffs assert that the defendants have copied a six-note musical sequence that is original to the plaintiff. First, we would need to know how likely it is for that six-note sequence to appear in a song. Note that copyright's originality doctrine doesn't require that the answer be a single time, in the plaintiffs' song, for the plaintiff to have a valid copyright.[80] Originality isn't novelty, and the plaintiffs may assert that the sequence is original to them even though it has appeared in other works.[81] Consider, also, that the estimate should reflect, to some degree, the rarity or commonality of the sequence in the relevant genre. These estimates are already incredibly difficult to produce, and they are, by far, the easier ones.

Next, the expert would have to determine the likelihood that the six-note sequence would appear in the defendant's song due to copying rather than independent creation. Consider a recent paper by Karol Jan Borowiecki that attempts to track musical influence over a five-century period of classical music.[82] The study creates similarity comparisons of compositions and composers using data from "18,074 melodic themes from 6,352 classical works by over 750 composers."[83] Using biographical and geographic data on

---

number of songs is, of course, increasing far more rapidly than the number of people. We thank Jamie Boyle and Jennifer Jenkins for this insight.

80.   *See* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345–46 (1991) ("Originality does not signify novelty; a work may be original even though it closely resembles other works so long as the similarity is fortuitous, not the result of copying. To illustrate, assume that two poets, each ignorant of the other, compose identical poems. Neither work is novel, yet both are original and, hence, copyrightable.").

81.   *Cf.* Sheldon v. Metro–Goldwyn Pictures Corp., 81 F.2d 49, 54 (2d Cir. 1936) ("[I]f by some magic a man who had never known it were to compose anew Keats's Ode on a Grecian Urn, he would be an 'author,' and, if he copyrighted it, others might not copy that poem, though they might of course copy Keats's."), *aff'd*, 309 U.S. 390 (1940).

82.   Karol Jan Borowiecki, *Good Reverberations? Teacher Influence in Music Composition Since 1450*, 130 J. POL. ECON. 991, 1077 (2022).

83.   *Id.* at 993.

composers' lives, the study compares the degree of similarity between teacher-student pairs with similarity between composers who were not in such relationships.[84] It finds that students' works tend to be more similar to their teachers' works than otherwise.[85]

This is some effort to get a sense of the degree of potential copying between musicians. Here, the presumption is that similarities that arise among student-teacher relationships are likely to be the result of copying, conscious or otherwise.[86] In some potential comparisons, the likelihood of copying given a degree of similarity is presumably very low, perhaps because the works arose fairly contemporaneously but in geographically disparate places. In these cases, access is unlikely. However, given the possibility of access, it gets much harder to estimate the actual likelihood of copying.

While this heroic effort tells us much about the possibility of musical copying over a long period of time and many musical compositions, the scale of the problem with contemporary music dwarfs these numbers. The study compared a mere 6,352 works, while Spotify and other streaming services currently receive as many as 120,000 new songs *every day*.[87]

One could plausibly use powerful digital technology to estimate the rarity of various sequences in the corpus of music uploaded to Spotify or Apple Music. With this data, experts could begin to determine whether two songs having a particular degree of similarity is rare or commonplace. This inquiry would need to incorporate the well-understood fact that a given note is more likely to be followed by some notes rather than by others. This decreases the total number of likely sequences, though not the total number of possible sequences, and increases the chance that existing similarities come from noncopying.

It is important to keep in mind that the experts in music copyright cases are attempting to testify that it's the amount of similarity—not the particular sequence at issue—that proves that copying occurred. As noted above—but easily missed by nonstatisticians—the opinion that similarities are so unusual that

---

84. *Id.* at 1004, 1010.

85. *See id.* at 1010, 1027 (stating the work of the student is more similar to the work of his teacher than other potential teachers).

86. *See id.* at 1014, 1027–28 (explaining that the students are imitating their teachers).

87. Murray Stassen, *There Are Now 120,000 New Tracks Hitting Music Streaming Services Each Day*, MUSIC BUS. WORLDWIDE (May 25, 2023), https://www.musicbusinessworl dwide.com/there-are-now-120000-new-tracks-hitting-music-streaming-services-each-day/ [http s://perma.cc/WPX7-YHRG].

they must result from copying is not an opinion about the sequence at issue; it is an opinion about how sequences come to be similar. Thus, we can't just ask about the sequence at issue, we would need to know how common it is for sequences of this type (properly defined) to be the same in two songs. That is, we would need to identify metrics for determining other sequences in songs that are as similar as the sequences of interest in the songs in suit.[88]

We might be able to measure note distances and lengths to identify comparable sequences, keeping in mind that copyright infringement claims in music often appeal to elements other than shared note sequences, including rhythm and timbre.[89] What we would end up with, if we could collect the information, is a dataset of songs and sequences within songs that could allow us to estimate how common it is for songs to share, say, a six-note pitch sequence. That would be part one.

Part two of the inquiry would be to determine how often such similarities result from independent creation (including use of public domain predecessors)[90] and how often they come from copying protected songs. Extremely rough estimates could come from various sources. One would be the presence of interpolation credits, which indicate that one song copied from another[91]— though this is complicated by the fact that such credits are sometimes given to settle disputes even when the songwriter denies copying.[92] As the caveat indicates, the assessment is not a fixed target—the rate of "copying" is dependent on what the law is, and changing practices about who gets interpolation credits would change our assessments of copying-based similarity going

---

88.   See discussion in Buccafusco, *supra* note 3, at 1648–49.

89.   *See, e.g.*, Structured Asset Sales, LLC v. Sheeran, No. 1:18-CV-05839, 2023 WL 3475524, at *2–3 (S.D.N.Y. May 16, 2023) (discussing combination of chord progression and harmonic rhythm and discussing other cases involving combinations of elements).

90.   Because everyone is free to draw on public domain predecessors, doing so counts as independent creation for purposes of copyright law. Buccafusco, *supra* note 3, at 1626–27.

91.   *See* Liesl Alyse Eschbach, *Do You Hear What I Hear?: The Inequities in Substantial Similarity Tests for Musical Copyright Infringement Cases*, 11 BERKELEY J. ENT. & SPORTS L. 71, 100 (2022) (defining interpolation as "the borrowing of pre-existing musical material and improvising to create a new work").

92.   *See* Jem Aswad, *Olivia Rodrigo Adds Paramore to Songwriting Credits on 'Good 4 U'*, VARIETY (Aug. 25, 2021, 7:38 AM), https://variety.com/2021/music/news/olivia-rodrigo-paramore-good-4-u-misery-business-1235048791/ [https://perma.cc/9S6V-HSRF] (noting "[r]etroactively-added songwriting credits have become increasingly common in recent years" due to prominent copyright infringement cases); Emma Perot, Music Copyright Ownership: Factors Behind the Surge in Writer Credits and Rights Clearance 19 (2023) (unpublished manuscript) (on file with authors).

forward.[93] We could also look for how many songs are explicitly labeled as arrangements. On the other side, we could attempt to identify how many sequences are found in public domain sources by using a date cutoff. This discussion also makes clear that base rates can be expected to vary across genres. Many people will announce their appropriation-based visual art;[94] this seems somewhat less likely with textual works.

Once we had some estimates of how often similarities result from copying and how often from noncopying, we could then start to make statements about the likelihood that a given instance of similarity is the result of copying. But we caution that the game might not be worth the candle: in most cases of putatively infringing similarity, similarity can make copying more likely than it would be in the absence of similarity, but it can't do much to tell us whether this similarity was the result of copying.[95]

Perhaps most fundamentally, all these efforts fail to grapple with estimating the effects of psychological processes in the defendant's mind.[96] Assuming the defendant actually heard the plaintiff's song, did the memory encode properly at the time? If it did encode, did the memory of the song degrade during the months or years between access and the alleged copying? And finally, even if there was a lingering, undegraded memory trace in the defendant's mind, did she *actually use* that memory trace when composing her own song? Experts cannot even begin to opine on any of these questions, and we doubt that they will ever be able to.[97]

93. *See* Perot, *supra* note 92, at 19, 21, 28 (explaining that copyright risk-aversion has combined with other changes in musical practice to substantially increase the number of credited writers in popular music).

94. *See* Blanch v. Koons, 467 F.3d 244, 247 (2d Cir. 2006) (stating Koons aimed to comment on mainstream media's influence on popular appetites through his depictions of women's feet in front of popular treats); *see also* Graham v. Prince, 265 F. Supp. 3d 366, 372–73 (S.D.N.Y. 2017) (explaining Prince screenshotted an Instagram post, which he then printed onto a canvas).

95. An additional problem is that our instances of copying-based and non-copying-based similarity may not be randomly selected, which can make it improper to extrapolate to other situations. Pharmacia Corp. v. Alcon Lab'ys, Inc., 201 F. Supp. 2d 335, 359–60, 362 (D.N.J. 2002). In substantial similarity cases, the people we have chosen to "test"—the defendants—are not randomly selected, but they are not selected from a group known to be at higher risk of copying either. They are selected instead from the more successful songwriters because that's who it makes sense to sue. Thus, we are not even as confident of their characteristics as we are of those of the overall group (because they are likely to diverge in some significant ways from the mean, median, or modal songwriter, including perhaps in the capacity to innovate and come up with unusual combinations of notes).

96. *See* Buccafusco, *supra* note 3, at 1647–49.

97. *Id.* at 1648–49.

It is important to note that our examples are from music, but the phenomenon also occurs in other creative fields. The problem of similar creations by different authors is often handled by the doctrine of scènes à faire, in which courts refuse to protect tropes and other creative elements that are standard for a genre.[98] But convention-influenced coincidence also occurs with elements that don't rise to the level of scènes à faire. Given how many books are written, for example, it is not shocking that two different fantasy authors writing in English independently chose similar names for their white "everyboy" protagonists: Larry Potter and Harry Potter.

IV. CURRENT EXPERT EVIDENCE ABOUT THE PROBABILITY OF COPYING IS UNRELIABLE AND SHOULDN'T BE ADMITTED

Given the current state of musicology, it seems impossible for expert witnesses to provide the appropriate data about the probability of similarities arising because of copying and the probability of similarities arising because of independent creation. Even the more sophisticated examples described in Part I fall far short of this sort of precision, especially when we realize how precise these numbers must be.[99] Recall that a less than 1% change in the base rate in the flour barrel example radically changed the inferences to be drawn from the facts.[100]

So, what should happen now? Under the new guidelines for F.R.E. 702, district court judges should prevent experts from testifying on the probability of copying in copyright infringement cases.[101] In recent years, scholars and commentors have grown concerned about insufficient oversight of federal judges concerning the reliability of expert testimony.[102] They worry that, in many forensic cases, experts have been overstating the reliability of their

---

98.   A.A. Hoehling v. Universal City Studios, Inc., 618 F.2d 972, 979 (2d Cir. 1980).

99.   *See supra* Section II.B.

100.   *See supra* notes 56–62 and accompanying text.

101.   *See* Memorandum from Hon. Patrick J. Schiltz, Chair, Advisory Comm. on Evidence Rules, to Hon. John D. Bates, Chair, Standing Comm. on Rules of Prac. & Proc. of the Jud. Conf. of the U.S. (May 15, 2022) [hereinafter Advisory Comm. Report], *available at* https://www.uscourts.gov/rules-policies/archives/committee-reports/advisory-committee-evidence-rules-may-2022 [https://perma.cc/GF6U-BJKR]; *supra* Section II.B. (discussing expert testimony regarding probability of copying without meaningful quantification, which is necessary for a reliable methodology to support expert testimony under the amended guidelines).

102.   *See, e.g.*, David E. Bernstein & Eric G. Lasker, *Defending* Daubert*: It's Time to Amend Federal Rule of Evidence 702*, 57 WM. & MARY L. REV. 1, 32, 35–36, 43–44 (2015).

methods, thereby leading juries to false inferences.[103] Recently, the Advisory Committee on the Rules of Evidence unanimously approved amendments to F.R.E. 702 to clarify and strengthen the requirement that expert testimony be based on reliable methods.[104]

The amendments emphasize that proponents of expert evidence must establish its reliability by a preponderance of the evidence.[105] Under the new amendments, reliability is clearly a matter of admissibility rather than merely of weight.[106] That is, if expert evidence isn't likely to be reliable, it shouldn't be admitted. The trial judge must determine whether "the expert's opinion reflects a reliable application of the principles and methods to the facts of the case."[107]

In light of this clarification, we think that district courts should prohibit musicologists from testifying about the probability of copying versus independent creation unless they can show that they have a mathematically reliable way of estimating those probabilities. And given our arguments above, we do not believe that they do. That doesn't mean that expert musicologists would have no role to play in copyright cases. Experts might helpfully point out the nature of the similarities and differences between the two works. They might also testify about the relative rarity or commonality of the musical components that are claimed to be copied. But they have no reliable means of estimating the probability of copying in light of this information, and they should be prevented from testifying to that end. Thus, even when experts are allowed to testify about musical tropes and variations, they should not be allowed to present the kinds of conclusions we saw in the examples above—that the similarity between two songs shows copying, plagiarism, infringement, or the like.

Defenders of this expert testimony might respond that similarity can be circumstantial evidence of copying, like a trout in the milk.[108] But the weight of circumstantial evidence depends

---

103.   *See id.* at 17, 45–46; Advisory Comm. Report, *supra* note 101.

104.   Advisory Comm. Report, *supra* note 101.

105.   *Id.*

106.   This has been the case since the *Daubert* decision, but courts have not consistently applied it. *Id.*

107.   *Id.*

108.   Shyamkrishna Balganesh and Peter Menell have made a version of this argument. Shyamkrishna Balganesh & Peter S. Menell, *Proving Copying*, 64 WM. & MARY L. REV. 299, 310 (2022) (defending their conception of the "inverse ratio" rule). However, as

on the tightness of the link between the evidence and the fact it is supposed to prove.[109] There are vanishingly few ways a trout can get in the milk without some kind of misfeasance. By contrast, in every copyright infringement case, there is at least one plausible way that the expression at issue could have been created without infringing: "independently," from the mind of the putative author. How do we know that? Well, because the author who created the plaintiff's work did it! That independent creation is the source of the plaintiff's ownership claim; without it, the plaintiff cannot bring an infringement claim. If it's possible for one author to independently conceive new expression, it's possible for another to do so.[110]

Of course, the more expression is at issue, the more unlikely independent creation seems, but it's essentially unheard of for defendants in wholesale copying cases to deny that there was copying (they're much more likely to deny that they were the downloaders than that the download was copied from a plaintiff's work). The expert testimony with which we are concerned is instead offered in cases of partial, fragmentary similarity, where it can have an outsized effect on factfinders because of the perceived weight of expertise and factfinders' general unfamiliarity with the details of creative production. But it is exactly in the cases involving fragmented similarity that we lack the relevant data, whether that's about a supposedly unique series of notes or a supposedly unique combination of individual features. And when we do have the data—when the plaintiff based its work on a specific public domain precedent—courts easily understand that they need to work harder to identify whether the defendant copied something protectable from the plaintiff or only copied expression from the predecessor work.[111]

---

they acknowledge, the weight of such evidence depends on knowing the underlying probabilities. *Id.* at 320. And that is precisely what we don't know (and they understandably don't attempt to define because that is not part of their theoretical project). They attempt to deal with the problem of base rates by arguing that when similarity is close to 100%, we don't need to worry much about probabilities. *See id.* at 328–29. But, if a situation requires an expert to detect the similarity, it is simply not in the range that would allow us to avoid inquiring into base rates.

109.   Buccafusco, *supra* note 3, at 1632–33 (explaining there are some similarities that are more indicative of copying than others, such as the mannequin with a second hairline, while others, such as a toy pig with a curled tail, are not).

110.   *See id.* at 1626, 1629, 1636. On copyright law's divergent treatment of originality and copying, see *id.* at 1650–55.

111.   *See, e.g.*, L. Batlin & Son, Inc. v. Snyder, 536 F.2d 486, 488–89 (2d Cir. 1976); Bridgeman Art Libr. v. Corel Corp., 36 F. Supp. 2d 191, 196–97 (S.D.N.Y. 1999).

Relatedly, "substantial similarity" is not a trout. What constitutes enough similarity to infringe is a hotly contested and notoriously ill-defined concept. In essence, we don't know what a "trout" is for copyright purposes. If we identified a previously unknown chemical in a carton of milk, we wouldn't thereby know whether it came by way of a cow.

Given the evidentiary uncertainties that are inevitable in cases where copying is contested, the concept of "striking similarity" used to infer copying should, if used at all, be limited to cases of much larger scale, unfragmented similarity such as a full verse, rather than deployed to cover individual elements of a work. The reality of base rate ignorance makes it unwise to infer copying from similarity alone, no matter how tempting the plaintiff's narrative is. Along with limiting the admissible scope of expert testimony, then, such testimony about similarity shouldn't be allowed to substitute for real evidence of access—after all, no expert is needed for a jury to infer that, where two, thousand-word poems are identical, one was copied from the other. Of note, no expert was required to understand the significance of the duplicate hairline in the mannequin case discussed above.[112] The expert's testimony is offered only where the similarities are coupled with differences, and it is precisely in that situation that the base-rate problem is biggest, and testimony about the likelihood of copying is most likely to be unfounded and misleading.

V. CONCLUSION

In fragmentary copying cases, plaintiffs claim that they created their works *ex nihilo*, without copying the expression at issue. They also claim that the defendants did not do the same. Plaintiffs' experts often translate infringement claims into arguments that the similarities between the works are such that independent creation is unlikely or essentially impossible. In the absence of sufficient information about base rates of similarity and rates of copying and not-copying this testimony is unfounded. Musicologists are rarely statisticians, and they should not make probability claims unless and until they have relevant evidence supporting such claims. This would be a small but worthwhile move towards honesty about what can and can't be known about copying.

---

112.   *See supra* text accompanying notes 21–22.