The Evidentiary Use and Misuse of Forensic Musicology in Copyright Litigation

Alfred C. Yen[*]

TABLE OF CONTENTS

 Professor of Law and Dean's Distinguished Scholar, Boston College Law School. This research was generously supported by grants from Boston College Law School. Earlier versions of this Article were presented at the Art Law Works in Progress Conference, the Intellectual Property Scholars Conference, faculty colloquia at the University of Arizona Law School and Tulane Law School, and as the Distinguished Shidler Lecture at the University of Washington Law School. The author is deeply appreciative for the feedback received there and initial guidance in musicology from Rose Subotnick and Alan Williams. Joseph Liu, Jessica Silbey, and Katherine Leo helpfully commented on earlier drafts. The author would also like to thank Celine Desantis, Iris Ryou, Melody Mathewson, Gwendolyn Ljung-Barth, J.P. LaBarge, Isaac Pavlik, and Hyungee Kim for their research assistance.

INTRODUCTION

This Article examines the use of testimony from expert forensic musicologists in music copyright infringement cases. It concludes that courts presently misuse such testimony in ways that confuse juries and violate the Federal Rules of Evidence. This does not mean that forensic musicologists should not testify in music copyright cases. Musicologists have ample training and expertise to help juries identify formal similarities between musical works[1] and understand whether those similarities suggest copying.[2]

Such expertise does not mean, however, that musicologists should also, as they presently often do, opine that two works are substantially similar or that identified similarities prove that a defendant has (or has not) copied from the plaintiff's work. Granted, Federal Rule of Evidence 704 states that "[a]n opinion is not objectionable just because it embraces an ultimate issue."[3] Nevertheless, a proper understanding of copyright law and the methodological foundation for such testimony reveal that the evidentiary risks of such testimony outweigh its benefits.[4]

This matters because forensic musicology testimony has become a major feature of music copyright litigation.[5] The typical music infringement case depends on inferences drawn from partial or fragmented similarities between the works in question, often similarities involving only a few bars of music.[6] These moments of

[1] *See infra* note 228 and accompanying text.

[2] *See infra* notes 264–274 and accompanying text.

[3] FED. R. EVID. 704(a).

[4] *See infra* Part II.C.

[5] *See* KATHERINE M. LEO, FORENSIC MUSICOLOGY AND THE BLURRED LINES OF FEDERAL COPYRIGHT HISTORY 3 (2021) (noting that musical experts have been part of copyright litigation for nearly 200 years, but that recognition of forensic musicology as a discipline arose in the late 20th century); Edward Lee & Andrew Moshirnia, *Do Experts Matter? A Study of the Effect of Musicologist Testimony in Copyright Cases*, 2022 U. ILL. L. REV. 707, 710 (2022) (reporting that "forensic musicologists have gained in both importance and notoriety").

[6] *See, e.g.*, Gray v. Hudson, 28 F.4th 87, 92 (9th Cir. 2022) (describing claim of infringement based on similarity between ostinatos — a repeating musical figure — in two works); Griffin v. Sheeran, 351 F. Supp. 3d 492, 499-500 (S.D.N.Y. 2019) (describing copyright infringement claim concerning two works sharing partial similarities in chord

suspicious similarity raise the possibility of copyright infringement, but they do not establish it because coincidence remains a possible explanation.[7] And indeed, defendants routinely deny that they have copied from the plaintiff.[8]

Forensic musicologists become star witnesses because copying is an essential element of copyright infringement.[9] Merely identifying similarities between two works will not prove that the composer of one copied from the other because works share similarities for many reasons. Similarities may exist because musical works follow specific compositional conventions.[10] They may exist because both composers used material from a common source.[11] And importantly, similarities sometimes occur by coincidence.[12] Plaintiffs therefore use forensic musicologists to rule out innocent explanations for suspicious similarities. Not surprisingly, defendants respond by employing musicologists to take the opposite position.

Allowing such testimony arguably makes sense. Music is a specialized field, so presumably the opinions of musicologists would help juries deliberate about infringement more effectively. However, closer examination shows that although musicological testimony can help juries, courts give such testimony more prominence than they should.

Under clearly established doctrine, a copyright plaintiff must prove two things to establish copyright infringement. First, the plaintiff must show that the defendant copied from the plaintiff's work.[13] Second, the

progressions, rhythm, and melody); Skidmore v. Led Zeppelin, 905 F.3d 1116, 1122 (9th Cir. 2018) (describing claim of infringement based on similarity between opening notes of two works).

[7] *See* J. Peter Burkholder, *Musical Borrowing or Curious Coincidence?: Testing the Evidence*, 35 J. MUSICOLOGY 223, 223-27 (2018) [hereinafter Burkholder, *Musical Borrowing*] (describing the musicological challenge of distinguishing musical similarities explained by copying from those explained by coincidence).

[8] *See supra* note 6 (listing cases in which defendants accused of copyright infringement denied copying).

[9] *See infra* Part II.A.

[10] *See infra* notes 264–265 and accompanying text.

[11] *See infra* notes 326–328 and accompanying text.

[12] *See* Burkholder, *supra* note 7, at 224-27 (acknowledging the possibility of similarities arising from coincidence and specifically noting that relatively brief sequences of similarity are more likely to arise from coincidence).

[13] *See infra* note 42 and accompanying text.

plaintiff must prove improper appropriation — namely that the defendant's copying causes lay listeners to consider the overall aesthetic impression of the defendant's work wrongfully similar to the plaintiff's.[14] These elements define the evidentiary purposes for which expert musicology testimony might be offered.

Initially, a plaintiff could offer an expert who identifies similarities between the works at hand and opines that the similarities establish copying. Such testimony combines musical knowledge with quasi-probabilistic intuition.[15] Additionally, a plaintiff could offer expert testimony to prove improper appropriation by having the expert assert that the identified similarities render the defendant's work a near aesthetic equivalent of the plaintiff's work. Here, the opinion no longer rests on musical knowledge and probability. Instead, the expert makes an aesthetic determination, opining that the two works are so aesthetically similar that (assuming that copying has otherwise been proven) infringement exists.[16] Of course, defendants could offer opposing expert testimony.

With respect to these possibilities, modern copyright doctrine clearly states that expert testimony is appropriate only to establish copying.[17] Courts believe, with some justification, that experts can help juries determine if copying exists.[18] If nothing else, musicologists can identify similarities between two works that a lay jury may not perceive, and they may be able to explain to jurors why those similarities do or do not suggest copying. With respect to improper appropriation, however, courts consistently hold that experts cannot help juries decide if two

[14] *See infra* note 42 and accompanying text.

[15] *See* Positive Black Talk, Inc. v. Cash Money Records, 394 F.3d 357, 370 (5th Cir. 2004) (stating that an inference of copying may be supported by "any similarities . . . that, in the normal course of events, would not be expected to arise independently in the two works"); *see also infra* Part II.A (describing content of forensic musicology opinions).

[16] *See infra* notes 51–52 and accompanying text (describing the difference between similarities probative of copying and similarities probative of improper appropriation).

[17] *See infra* notes 256–259.

[18] *See* Burkholder, *Musical Borrowing*, *supra* note 7, at 224-27 (describing how musical analysis might support conclusions about the existence of copying); *infra* note 257 and accompanying text (discussing leading cases in which courts permit expert testimony to help juries determine if copying exists).

works share too much aesthetic similarity.[19] The test for improper appropriation depends on the perception of lay listeners,[20] and jurors are presumably lay listeners. By definition, expert musicologists are not lay listeners, making their aesthetic perceptions irrelevant to the jury's deliberation.[21]

Problems initially arise because courts do not effectively limit forensic musicologists to opinions about copying. This happens because incorrect and confusing elaborations of the law collapse the distinction between copying and improper appropriation by using "substantial" to describe the degree of similarity required to prove each element. Courts therefore mistakenly think that expert witnesses address copying when they are actually testifying about improper appropriation.[22] This leads to problems of prejudice and confusion under Federal Rule of Evidence 403.[23]

On those occasions where forensic musicologists address copying, alternate problems arise because courts seem unaware that the field of musicology does not offer musicologists the tools to determine accurately or consistently when copying exists. This causes the admission of expert testimony that is not sufficiently reliable to pass muster under Federal Rule of Evidence 702.

This second observation may seem puzzling because musicologists study copying.[24] One might therefore imagine that the field of musicology has developed methods that reliably identify when copying has occurred. Unfortunately, no such methods exist. Although musical analysis can identify clues about whether one composer has copied from another, it cannot determine when those clues amount to solid evidence of copying. Musicological conclusions about copying therefore depend heavily on the personally idiosyncratic intuitions and biases of the

[19] *See infra* notes 256–259.

[20] *See infra* note 50.

[21] *See infra* notes 258–260 and accompanying text.

[22] *See infra* note 276 and accompanying text.

[23] *See infra* notes 178–201, 274–280 and accompanying text (explaining Rule 403 and problems of confusion raised by improper use of forensic musicology testimony).

[24] *See* J. Peter Burkholder, *The Uses of Existing Music: Musical Borrowing as a Field*, 50 NOTES 851, 851 (1994) [hereinafter Burkholder, *Uses of Existing Music*] (stating that "[m]usicologists have studied musical borrowings for over a century").

musicologist. A forensic musicologist may sincerely believe that her conclusions about copying are correct, but that is not enough to make her opinion reliable under the Federal Rules of Evidence.

To the extent that forensic musicologists base their opinions on the logic of probability, they lack crucial information necessary to make sound statistical inferences. As Professors Christopher Buccafusco and Rebecca Tushnet have pointed out, statistical inferences about the likelihood of copying require information about the relative likelihood that observed similarities arise from copying or coincidence.[25] Forensic musicologists do not have this information, nor can they make up for it by opining that observed similarities are rare.[26]

If forensic musicologists claim that experience has taught them the difference between innocent, suspicious similarities and those caused by copying, it is difficult — if not impossible — to see how the study of music creates the foundation of reliability required by the Federal Rules of Evidence.[27] For example, one might believe that musicologists learn to reliably distinguish copying from coincidental similarity by studying numerous exemplars of these phenomena. Unfortunately, it is effectively impossible for a musicologist to find these hypothesized exemplars because composers who create suspicious similarities do not verify how the similarities arose.[28]

Finally, if forensic musicologists assert that the field of musicology has developed methods of musical analysis that reliably lead to objective conclusions about copying, an examination of musicological argument reveals no such methods. Indeed, musicologists typically "prove" copying by identifying similarities between works and asserting that copying exists without referencing any accepted or tested method of proof. Musicologists may identify the same similarities, and they may agree on how to describe them. However, nothing tells a musicologist when enough similarities exist to establish copying. Accordingly, a

[25] Christopher Buccafusco & Rebecca Tushnet, *Of Bass Notes and Base Rates: Avoiding Mistaken Inferences About Copying*, 61 HOUS. L. REV. 235, 246-53 (2023).

[26] *Id.*

[27] *See* FED. R. EVID. 702 (permitting qualification of expert witness on the basis of experience).

[28] *See infra* note 295 and accompanying text.

musicologist's conclusion of copying frequently depends heavily, if not entirely, on personal fiat.[29]

These issues make the overreliance on forensic musicology deeply problematic under the Federal Rules of Evidence. Rule 702 permits expert testimony only if the expert's testimony is based on sufficient facts or data, the testimony is the product of reliable principles or methods, and the expert has reliably applied the principles and methods to the facts of the case.[30] These criteria separate inadmissible speculation and personal hunch from admissible opinions based on reliable knowledge, experience, and method.[31] Unfortunately, it is likely that opinions about copying from forensic musicologists embody the kinds of speculation and personal hunch that the Federal Rules of Evidence exclude. Courts should therefore revise how forensic musicologists testify to ensure that juries hear only helpful and reliable evidence.

For example, courts should permit musicologists to testify about things over which sufficient musicological consensus exists to ensure reliability, such as musical terminology and the identification of formal similarities between two musical works.[32] Courts should also generally permit testimony about musical conventions and the occurrence of specific musical features in other works to help juries understand how certain similarities are or are not probative of copying.[33] Courts should not, however, allow experts to opine that two works are substantially similar or offer opinions about whether copying occurred because these

[29] *See infra* notes 299–362 and accompanying text.

[30] FED. R. EVID. 702.

[31] *See* Kumho Tire Co. v. Carmichael, 526 U.S. 137, 157 (1999) (citing Gen. Elec. Co v. Joiner, 522 U.S. 136 (1997)) (affirming exclusion of expert testimony that resting too heavily on expert's personal opinion); *Joiner*, 522 U.S. at 140-41 (upholding exclusion of medical causation testimony characterized by the district court as "subjective belief or unsupported speculation"); United States v. Frazier, 387 F.3d 1244, 1272 (11th Cir. 2004) (Federal Rules of Evidence ensure that juries rely on "relevant and credible evidence, rather than on speculation or otherwise unreliable conjecture"); Oddi v. Ford Motor Co., 234 F.3d 136, 158 (3d Cir. 2000) (stating that expert's opinion must not be based on subjective belief or unsupported speculation).

[32] *See infra* note 263 and accompanying text.

[33] *See infra* notes 264–272 and accompanying text.

opinions will be unreliable and confusing to the jury.[34] Musicologists may render these opinions during the academic study of music, but that does not mean that the opinions meet our legal system's standards for admissible evidence.[35]

The pages that follow elaborate the foregoing. Part I reviews the law of copyright infringement. After laying out the basics, the Part describes how courts have inaccurately elaborated the law. Part II describes the substantive content of forensic musicology testimony commonly presented during infringement litigation, and it reviews basic principles of evidence law and their application to forensic musicology testimony in music infringement cases. It describes how, under existing practice, courts allow forensic musicologists to testify on matters previously deemed inappropriate for expert testimony, thereby misleading the jury. It also examines the field of musicology to show that a musicologist's opinions about the existence of copying lack sufficient reliability to pass muster under the Federal Rules of Evidence. Part III builds on Part II to recommend how courts can ensure the proper use of forensic musicology to make music infringement cases less confusing to juries. Part IV considers some possible objections to the proposals made here, and Part V concludes.

I. PROVING MUSIC COPYRIGHT INFRINGEMENT

The basic elements of copyright infringement, along with their confused judicial elaboration, help explain the proper and improper uses of forensic musicology opinions.

---

[34] *See infra* notes 276–376 and accompanying text.

[35] A similar point is made by the Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, in Strengthening Forensic Science in the United States: A Path Forward. NATIONAL RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 15 (distinguishing legal factfinding from scientific inquiry and noting "fundamental limitations" in forensic science disciplines). That report identified "a serious problem" arising from insufficient validation of methods relied upon by various forensic sciences. *Id.* at 8. The report states that courts have not consistently or clearly ensured the reliability of forensic testimony. *Id.* at 11. The report also states that more support and development must be given "to all credible forensic science disciplines if they are to achieve the degrees of reliability needed to serve the goals of justice." *Id.* at 13.

A. The Basic Framework: Copying and Improper Appropriation

Arnstein v. Porter[36] provides the framework that governs proof of copyright infringement. The case involved a claim of music copyright infringement brought by Ira Arnstein against Cole Porter. Arnstein's claims struck the trial judge as "fantastic"[37] for good reason. Arnstein apparently testified in his deposition that Porter had hired "stooges" to burglarize his apartment and steal his music.[38] Porter, of course, denied this, and Arnstein admitted he had nothing more than suspicion to link Porter to the alleged burglaries.[39] On these facts, the district court understandably granted summary judgment for Porter.[40] The Second Circuit reversed, stating that the district court should have allowed Arnstein to argue his case to the jury, however improbable his allegations might have been.[41] In so doing, the Second Circuit provided a simple, two-element framework for proof of infringement. First, the plaintiff must prove that the defendant copied from the plaintiff's work. Second, the plaintiff must show that the defendant's copying amounted to "improper appropriation" — namely copying that rises to the level of infringement.[42]

The Second Circuit then offered guidance about establishing each of these elements. For copying, a plaintiff could rely on direct evidence

[36] 154 F.2d 464 (2d Cir. 1946).

[37] *Id.* at 469.

[38] *Id.* at 467; *see also* GARY A. ROSEN, UNFAIR TO GENIUS: THE STRANGE AND LITIGIOUS CAREER OF IRA B. ARNSTEIN 221 (2012) (describing Arnstein as a "crank" and noting Arnstein's allegations about how "stooges" hired by Porter broke into Arnstein's apartment to steal his works).

[39] *Arnstein*, 154 F.2d at 467.

[40] *Id.* at 468.

[41] *Id.*

[42] *Id.* Many courts follow this two-step framework, even if their doctrinal statements lack the clarity of *Arnstein*. *See* Design Basics, LLC v. Signature Const., Inc., 994 F.3d 879, 887 (7th Cir. 2021) (noting that infringement requires proof of copying and improper appropriation); Compulife Software, Inc. v. Newman, 959 F.3d 1288, 1301-02 (11th Cir. 2020) (referring to "factual" copying and "actionable" copying); Batiste v. Lewis, 976 F.3d 493, 501-02 (5th Cir. 2020) (stating that plaintiff must show that the defendant "actually used" material from plaintiff's work and that the copying went so far as to constitute infringement); Cortés-Ramos v. Martin-Morales, 956 F.3d 36, 41 (1st Cir. 2020) (holding that a plaintiff must prove actual copying and a degree of copying that constitutes infringement).

such as a defendant's admission of copying.[43] Alternatively, a plaintiff could build a circumstantial case by proving that the defendant had access to the plaintiff's work and that similarities between the defendant's work and the plaintiff's work demonstrated copying.[44] In unusual cases, the two works might be so similar ("strikingly similar") that independent evidence of access would not be necessary.[45] The logic for this is probabilistic. Works often share similarities that do not support inferences of copying. Two paintings may feature the Eiffel Tower and Seine River, but these landmarks appear in paintings so often that copying provides an implausible explanation for the similarity. By contrast, if a detective novel contains the words "'Twas brillig, and the slithy toves,"[46] an inference of copying becomes quite strong because the words are uniquely associated with Lewis Carrol's well-known poem "Jabberwocky." It is therefore unlikely that the detective novel's author would have independently created the same words.

For improper appropriation, the court stated that the test "is the response of the ordinary lay hearer."[47] This meant playing the two works so that the jurors (presumably as lay listeners) could determine if the defendant's work was "illicitly" or "improperly" similar to the plaintiff's.[48] The reasoning here is purely aesthetic, not probabilistic. In the court's mind, copyright infringement exists only if the defendant's copying resulted in a work whose aesthetic appeal was sufficiently close to that of the plaintiff's that a jury would find it wrongful.[49]

The foregoing shows that proof of infringement requires the identification and appraisal of similarity in two distinct ways. With respect to copying, one must identify similarities and ask whether those similarities support an inference that the defendant has copied from the

---

[43] *Arnstein*, 154 F.2d at 468.

[44] *Id.*

[45] *Id.*

[46] LEWIS CARROLL, THROUGH THE LOOKING GLASS AND WHAT ALICE FOUND THERE 19 (1882).

[47] *Arnstein*, 154 F.2d at 468.

[48] *Id.* at 473.

[49] *Id.* ("The question, therefore, is whether defendant took from plaintiff's works so much of what is pleasing to the ears of lay listeners, who comprise the audience for whom such popular music is composed, that defendant wrongfully appropriated something which belongs to the plaintiff.").

plaintiff's work. This means identifying similarities between two works and asking whether the features in question are relatively common or rare. If they are common, the inference of copying weakens because it implies that many composers — including the defendant — have independently created the same features. Conversely, if the features in question are rare, the inference of copying is strengthened.

For improper appropriation, one must identify similarities and ask how those similarities affect the aesthetic impression made by the works in question. As *Arnstein* stated, the determination at hand is "whether defendant took from plaintiff's works so much of what is pleasing to the ears of lay listeners, who comprise the audience for whom such popular music is composed, that defendant wrongfully appropriated something which belongs to the plaintiff."[50] Crucially, this determination does not depend on the frequency or rarity of features shared by two works. What matters instead is the way the similarities affect the overall aesthetic impression the works make.

For purposes of this Article, I will call similarities supporting inferences of copying "probative similarities" and similarities supporting improper appropriation "aesthetic similarities."[51] Similarities between two works can be simultaneously probative and aesthetic. If two songs share the same melody, this similarity suggests copying and makes the works sound alike. This does not mean, however, that probative similarities establish improper appropriation or that aesthetic similarities establish copying.[52]

One can be quite confident that a defendant has copied from another with no effect on overall aesthetic similarity. Consider again our

[50] *Id.*

[51] The term "probative similarity" first appeared in Alan Latman's influential article about copyright infringement. Alan Latman, *Probative Similarity as Proof of Copying: Toward Dispelling Some Myths in Copyright Infringement,* 90 COLUM. L. REV. 1187, 1190 (1990) (suggesting use of the term "probative similarity"). Professor Latman's article is the leading article identifying and explaining the distinction under discussion here and the problems courts have had keeping the distinction between probative and aesthetic similarities straight.

[52] *Id.* at 1214; *see also* Zalewski v. Cicero Builder Dev. Inc., 754 F.3d 95, 101 (2d Cir. 2014) (noting that similarity between two works is often relevant to both copying and improper appropriation, but also noting that similarity sometimes supports only one or the other).

hypothetical detective novel containing the words "'Twas brillig, and the slithy toves." As noted earlier, the unusual nature of these words creates a strong inference of probative similarity.[53] However, this probative similarity creates no meaningful inference of aesthetic similarity. The overall appeal of a detective novel surely differs from that of Carroll's poem.[54]

Similarly, one can perceive strong aesthetic similarity that creates no inference of copying. For example, two musical works may sound similar because they share the features of a rock beat, male vocals, drums, and an electric guitar solo.[55] Blues songs typically employ a specific harmonic plan that occurs over twelve bars.[56] Mariachi music typically employs a distinctive combination of violins, guitars, and trumpets.[57] Furthermore, composers and performers often make use of motifs and riffs that, while perhaps not specific to a particular genre's convention, have become so common that they have been integrated into the "language" of music.[58] The common existence of these features may be

[53] *See supra* note 46 and accompanying text.

[54] To provide one musical example of this, consider Debussy's "Golliwogg's Cakewalk," which is widely considered to have a borrowing from Wagner's *Tristan and Isolde*. However, Debussy's use of the material is generally considered humorous, giving his work an extremely different aesthetic impression. *See* Mark DeVoto, *The Strategic Half-Diminished Seventh Chord and the Emblematic Tristan Chord: A Survey from Beethoven to Berg*, 4 INT'L J. MUSICOLOGY 139, 146 (1995) (noting "several writers" have noted the apparent borrowing by Debussy); Gregory Marion, *Crossing the Rubicon: Debussy and the Eternal Present of the Past,* 27 INTERSECTIONS 26, 41 (2007) (referring to "reference" by Debussy and characterizing it as parody); Burkholder, *supra* note 7, at 229-30 (referring to such borrowing).

[55] *See* Tisi v. Patrick, 97 F. Supp. 2d 539, 543 (S.D.N.Y. 2000) (noting that "much of rock music sounds the same" and attributing similarities between plaintiff and defendant's works to features of "standard usage in rock music").

[56] *See* JANE PIPER CLENDINNING & ELIZABETH WEST MARVIN, THE MUSICIAN'S GUIDE TO THEORY AND ANALYSIS 708-10 (4th ed. 2021) (describing basic aspects of the twelve-bar blues); *Understanding the 12-Bar Blues*, PBS, https://www.pbs.org/theblues/classroom/essays12bar.html (last visited Sept. 20, 2024) [https://perma.cc/W49D-SWTF].

[57] *See Mariachi*, ENCYC. BRITTANICA, https://www.britannica.com/art/mariachi (last visited July 9, 2024) [https://perma.cc/7EUU-J7N5].

[58] *See* Burkholder, *Musical Borrowing*, *supra* note 7, at 234 (referring to "shared conventions of the musical language").

musically significant and contribute a great deal to an impression of musical similarity, but it creates no reasonable inference of copying.[59]

Many cases confirm the *Arnstein* distinction between probative similarity and aesthetic similarity, making the importance of their independent proof clear.[60] Courts should not conclude that infringement exists simply because two musical works share many similarities. Instead, courts must determine whether enough probative similarities exist to establish copying and then measure whether the works share enough aesthetic similarity to support a conclusion of improper appropriation.

B. Confusion in the Law of Copyright Infringement

The two-element framework established by *Arnstein* tells us how to analyze a copyright infringement claim, but it does not tell us how much similarity it takes to establish copying or improper appropriation. We know that, assuming the defendant had access to the plaintiff's work, a court should find that the defendant has copied from the plaintiff if there are enough probative similarities between the two works to make copying more likely than independent creation or origination from a

[59] *See id.*

[60] *See* Design Basics, LLC v. Signature Const., Inc., 994 F.3d 879, 888 (7th Cir. 2021) (distinguishing between the kinds similarity required to prove copying and improper appropriation, and using the term "substantial similarity" only for purposes of improper appropriation); Rentmeester v. Nike, Inc., 883 F.3d 1111, 1117 (9th Cir. 2018) (requiring copying and "unlawful appropriation" for proof of infringement, and noting that the same term "substantial similarity" has been used to describe the level of similarity needed to establish copying and improper appropriation); Zalewski v. Cicero Builder Dev., Inc., 754 F.3d 95, 100 (2d Cir. 2014) (citing Arnstein v. Porter, 154 F.2d 464 (2d Cir. 1946)) (emphasizing the distinction between copying and wrongful copying); Johnson v. Gordon, 409 F.3d 12, 18 (1st Cir. 2005) ("The requirement of probative similarity is somewhat akin to, but different than, the requirement of substantial similarity that emerges at the second step [of infringement analysis]."); Bridgmon v. Array Sys. Corp., 325 F.3d 572, 576 n.7 (5th Cir. 2003) (using probative similarity and substantial similarity as "analytically distinct inquiries"); Dam Things from Denmark, a/k/a Troll Co. ApS, v. Russ Berrie & Co., 290 F.3d 548, 562 (3d Cir. 2002) (stating that probative similarity relates to "actual copying" and substantial similarity relates to improper appropriation); Repp v. Webber, 132 F.3d 882, 889 (2d Cir. 1997) (clearly identifying copying and improper appropriation as prerequisites to finding of infringement).

common source.[61] Similarly, a court should find improper appropriation when two works share so much aesthetic similarity that their overall impression is the same to a lay audience.[62] But how much probative similarity does it take to prove copying? And how much aesthetic similarity does it take for two works to share the same overall appeal to a lay audience?

The answers to these questions are effectively unknowable because they rest on fundamentally subjective human judgment. There is no precise, objective level of similarity that establishes copying, nor is there such a level of similarity that establishes improper appropriation. The trier of fact must rely on intuition to answer whether enough probative and aesthetic similarity exist to support copying and infringement.[63]

These challenges of subjectivity have greatly affected the development of copyright law, and not necessarily for the better. Judges want guidance about when findings of copying and improper appropriation are warranted, and they want to explain their decisions as consistent with the law. Courts have therefore tried to clarify the law of infringement by interpreting, elaborating, and reformulating the two-step *Arnstein* framework in hopes of making it more intelligible and concrete. Unfortunately, these efforts have not had the desired effect and have instead introduced confusion and error.[64] In fact, because courts have not paid sufficiently close attention to the distinction between probative similarity and aesthetic similarity, many opinions reach conclusions about infringement without clearly considering whether sufficient probative similarity exists to support a finding of

[61] *See supra* Part II.A.

[62] *See id.*

[63] *See* Nichols v. Universal Pictures Corp. 45 F.2d 119, 121 (2d Cir. 1930) (stating that "nobody has ever been able to fix" the boundary between permissible borrowing and actionable copying, "and nobody ever can").

[64] *See* Shyamkrishna Balganesh, *The Questionable Origins of the Copyright Infringement Analysis*, 68 STAN. L. REV. 791, 794 (2016) (describing modern infringement test as "a virtual black hole in copyright jurisprudence"); Mark A. Lemley, *Our Bizarre System for Proving Copyright Infringement*, 57 J. COPYRIGHT SOC'Y 719, 741 (2010) (stating that "[o]ur rules for proving copyright infringement make little sense"); Pamela Samuelson, *A Fresh Look at Tests for Nonliteral Copyright Infringement*, 107 NW. U. L. REV. 1821, 1823 (2013) (attributing difficulties in the analysis of copyright infringement to "too many tests and not enough guidance").

copying. This clearly violates the basic doctrine that copying is an essential element of copyright infringement, and it has created a pattern of governing precedent rife with inconsistency.

Two primary developments bear responsibility for this error and confusion. First, several courts, including the Second Circuit, have used the term "substantial similarity" to describe the kinds of similarity that establish copyright infringement, but they did not indicate clearly whether they were addressing probative or aesthetic similarity. Second, the Ninth Circuit has led many courts to exacerbate this problem by reformulating *Arnstein*'s two-part test in terms of "extrinsic" and "intrinsic" similarity.[65]

1. The Imprecise and Confusing Use of "Substantial Similarity"

Before *Arnstein,* many decisions used the terms "substantial" and "substantial similarity" to define whether a defendant's alleged copying rose to the level of copyright infringement. For example, in *Eggers v. Sun Sales Corp.*, the Second Circuit wrote that "where the subject of copyright is a book, infringement consists in the copying of some substantial and material part thereof."[66] Other courts used "substantial similarity" to explain findings about the existence of infringement.[67]

These uses of "substantial" and "substantial similarity" are entirely consistent with the principle that copying alone does not establish

[65] *See infra* notes 89–128 and accompanying text.

[66] Eggers v. Sun Sales Corp., 263 F. 373, 375 (2d Cir. 1920); *see also* Kustoff v. Chaplin, 120 F.2d 551, 560 (9th Cir. 1941) ("In order to prove infringement there must have been copying of a substantial portion of the copyrighted work."); Dymow v. Bolton, 11 F.2d 690, 691 (2d Cir. 1926) (copyright infringement exists only if defendant has copied "some substantial and material part" of plaintiff's copyrightable material); Wilson v. Haber Bros., 275 F. 346, 347 (2d Cir. 1921) ("Infringement of a copyright consists in the copying of some substantial and material part" of the plaintiff's copyrighted material); De Montijo v. 20th Century Fox Film Corp., 40 F. Supp. 133, 138 (S.D. Cal. 1941) (copyright infringement requires substantial reproduction of a copyrighted original); Frankel v. Irwin, 34 F.2d 142, 143 (S.D.N.Y. 1918) (infringement requires copying some substantial part of plaintiff's copyrighted material).

[67] *See* McMahon v. Harms, 42 F. Supp. 779, 780 (S.D.N.Y. 1942) (defendant's composition could not infringe because it lacked substantial similarity to plaintiff's); Arnstein v. Broad. Music, Inc., 46 F. Supp. 379, 381 (S.D.N.Y. 1942) (plaintiff's infringement claim failed for, among other things, failure to prove substantial similarity).

infringement. "Substantial similarity" is simply a way of stating that a defendant infringes when his copying results in a work that is too similar to the plaintiff's. Proof of copying remains separate from any determination or perception of "substantial similarity." Thus, proof of copying rests on probative similarity, while improper appropriation rests on the existence of substantial aesthetic similarity between the defendant's and plaintiff's work.

Yet not long after *Arnstein* was decided, courts began suggesting that "substantial similarity" mattered to the proof of copying. The 1958 case of *Costello v. Loew's, Inc.*[68] offers an early example. In that case, the District of D.C. Court used "substantial similarity" to apply the *Arnstein* test, but in doing so, it blurred the distinction between probative and aesthetic similarity. The *Costello* plaintiff alleged that the defendant's movie "Knights of the Round Table" infringed her drama "The Sangreal."[69] In deciding the case for the defendant on summary judgment, the court stated that infringement requires "(1) access; (2) substantial similarities between the two works; and (3) copying of the plaintiff's work by the defendant."[70]

This statement of the law does not match the two-step framework of *Arnstein*. *Arnstein* required proof of copying and improper appropriation, but *Costello* identified three elements of infringement: namely access, substantial similarity, and copying. It is possible to understand this elaboration of *Arnstein* as nothing more than an awkward restatement. After all, *Costello* identifies copying as a requirement, so "substantial similarity" could simply express the idea that improper appropriation occurs when the defendant's work bears too much aesthetic similarity to the plaintiff's, thereby leaving access as a superfluous separate requirement given its role in proving copying. Such a reading would continue properly limiting the significance of substantial similarity to improper appropriation.

However, the *Costello* court's application of its stated test complicates this neutral understanding. Because the defendant had admitted access,

[68] Costello v. Loew's, Inc., 159 F. Supp. 782 (D.D.C. 1958).

[69] *See id.* at 783.

[70] *Id.*

the court focused its attention on copying and improper appropriation.[71] The court first adopted the proposition that "for infringement or misappropriation there must be substantial similarities between the two works."[72] After that, the court stated that "[a]s to the element of copying . . . the appropriation by the defendant must have been of a substantial or material part of the protected work."[73] According to the court, copying may be inferred upon access and a sufficient degree of similarity, but "the similarity must be recognizable on ordinary observation."[74]

These statements show that the *Costello* court failed to maintain the distinction between probative and aesthetic similarity, opening the door to using "substantial" to describe the different similarities that establish copying and improper appropriation. The statement that "infringement or misappropriation" requires "substantial similarities between the two works" invokes improper appropriation.[75] However, by stating that copying requires appropriation of a "substantial or material" part of the plaintiff's work and that such similarity must be "recognizable on ordinary observation," the court has injected the ordinary observer from improper appropriation into the test for copying, and it has used the word "substantial" to describe the kind of borrowing that establishes copying.[76]

Unfortunately, the confusion exhibited by *Costello* appeared in other opinions about copyright infringement. Consider the 1966 case of *Ideal Toy Corp. v. Fab-Lu Ltd.*[77] in which the plaintiff claimed that the defendant's Randy and Mary Lou dolls infringed the plaintiff's Tammy

[71] *See id.* ("Inasmuch as the defendant admits corporate access to the plaintiff's work in 1934 and 1935, it is unnecessary to discuss the law concerning proof of that element.")

[72] *Id.*

[73] *Id.* at 784.

[74] *Id.*

[75] *Id.* at 783 (noting agreement of counsel that "for infringement of misappropriation there must be substantial similarities between the two works").

[76] *Id.* at 784 ("As to the element of copying . . . the appropriation by defendant must have been of a substantial or material part of the work . . . . [T]he similarity must be recognizable on ordinary observation.").

[77] 360 F.2d 1021 (2d Cir. 1966).

and Pepper dolls.[78] The district court issued a preliminary injunction in the plaintiff's favor, and the Second Circuit affirmed on appeal.[79]

The issuance of the preliminary injunction implied that the district court considered the plaintiff's case strong on both copying and improper appropriation. One would therefore expect a decision affirming the district court to consider both elements. This would require assessment of probative and aesthetic similarity. However, the Second Circuit did not conduct this analysis.

The trouble began with the court's apparent definition of copyright infringement. The opinion stated, "It is well established, however, that in order to sustain a claim of copyright infringement the claimant is required to demonstrate a substantial similarity between the copyrighted work and the alleged copy."[80] The court went on to state that substantial similarity exists when "an average lay observer would recognize the alleged copy as having been appropriated from the copyrighted work."[81]

This language clearly comes from the improper appropriation element of copyright infringement. However, if the Second Circuit were conducting a complete analysis of infringement, then the court clearly erred because it overlooked the element of copying. Even if one somehow accepted the court's definition of infringement as comprehensive, doing so would require combining copying and improper appropriation into a single analysis governed by substantial similarity. This means ignoring the distinction between probative similarity and aesthetic similarity and accepting one form of similarity as proof of the other.

In short, one cannot understand *Ideal Toy* as well-reasoned. The court either omitted a crucial element from its analysis, or it collapsed the *Arnstein* two-step framework in a way that does not make sense. Of these two possibilities, the former is more benign because it does not imply a change in the law. Perhaps *Ideal Toy* exhibited a simple error that other courts would not repeat.

[78] *See id.* at 1022.

[79] *See id.* at 1022-23.

[80] *Id.* at 1022.

[81] *Id.*

Unfortunately, other Second Circuit opinions also collapsed *Arnstein*'s two step framework, thereby deepening confusion and error in the law of infringement. For example, in *Reyher v. Children's Television Workshop*, the plaintiffs alleged that the defendants based their illustrated story *The Most Beautiful Woman in the World* on the plaintiffs' book entitled *My Mother Is the Most Beautiful Woman In The World*.[82] The district court dismissed the plaintiffs' case on the ground that there was "no substantial similarity between the two works as to copyrightable matter."[83]

In affirming, the Second Circuit stated that, "In an infringement action, a plaintiff must establish ownership of the copyright and copying by the defendant."[84] The court then wrote, "Because of the inherent difficulty in obtaining direct evidence of copying, it is usually proved by circumstantial evidence of access to the copyrighted work and substantial similarities as to protectible material in the two works."[85]

This reasoning duplicates *Ideal Toy*'s collapse of copying and improper appropriation into a single element governed by substantial similarity. To begin with, the court apparently eliminated improper appropriation as an element of infringement by stating that infringement requires only plaintiff's ownership of the copyright and copying. Next, the court explained that a plaintiff typically proves copying through access and "substantial similarities." Given the prior use of "substantial similarity" as a stand-in for the degree of aesthetic similarity needed to establish improper appropriation, the *Reyher* court all but explicitly stated that questions of copying rest on aesthetic similarity, not probative similarity.

These implications became explicit in *Warner Bros. Inc. v. American Broadcasting Cos. Inc.* where the Second Circuit once again stated that copying depends on a showing of substantial similarity.[86] The court wrote that the test for substantial similarity is "whether an average lay

[82] Reyher v. Child.'s Television Workshop, 533 F.2d 87, 89 (2d Cir. 1976).

[83] *Id.* at 88.

[84] *Id.* at 90.

[85] *Id.*

[86] Warner Bros. v. Am. Broad. Cos. Inc., 654 F.2d 204, 207 (2d Cir. 1981) ("[I]t is well settled that copying may be inferred where a plaintiff establishes that the defendant had access to the copyrighted work and that the two works are substantially similar.").

observer would recognize the alleged copy as having been appropriated from the copyrighted work."[87] Under *Arnstein*, this lay observer test should be used to assess whether improper appropriation exists. Its use to determine whether copying exists therefore confirmed what *Reyher* strongly implied, namely, that a test based on aesthetic similarity could be used to establish both copying and improper appropriation.[88]

Judicial inattention to the distinction between probative and aesthetic similarity greatly complicates the law of copyright infringement. Without a clear distinction between probative and aesthetic similarity, it is impossible to ensure that decisions about copyright infringement are properly reasoned because courts can find infringement without analyzing whether a defendant copied from the plaintiff's work. Unfortunately, this problem has worsened because the Ninth Circuit attempted to clarify the law by reformulating the two step *Arnstein* framework in a manner that adds even more confusion to the law of copyright infringement.

2. The Extrinsic/Intrinsic Analysis of the 9th Circuit

The Ninth Circuit case *Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp.*[89] rivals *Arnstein v. Porter* for influence over the law of copyright infringement. However, unlike *Arnstein*, *Krofft* did not provide a sensible framework to govern the proof of copyright infringement. Instead, *Krofft* offered a new two-step framework that fundamentally

[87] *Id.* at 208.

[88] Not surprisingly, other opinions exhibit the same confusion. *See* Craft Smith, LLC v. EC Design, LLC, 969 F.3d 1092, 1101-04 (10th Cir. 2020) (using "substantial similarity" to evaluate copying and improper appropriation); Enchant Christmas Light Maze & Mkt. Ltd. v. Glowco, LLC, 958 F.3d 532, 538-39 (6th Cir. 2020) (same); Blehm v. Jacobs, 702 F.3d 1193, 1199-1204 (10th Cir. 2012) (same); *see also* Amy B. Cohen, *Masking Copyright Decisionmaking: The Meaninglessness of Substantial Similarity*, 20 UC DAVIS. L. REV. 719, 735-39 (1987) (criticizing Second Circuit cases for failing to distinguish copying from improper appropriation); Lemley, *supra* note 64, at 719-20 (noting the difference between two kinds of similarity and lamenting that "[u]nfortunately courts and commentators regularly lump the entire inquiry under the single term 'substantial similarity'").

[89] 562 F.2d 1157 (9th Cir. 1977).

ignored *Arnstein*'s logic. This was a clear error that caused great confusion in the law.[90]

In *Krofft*, the plaintiffs were two corporate entities, apparently named after Sid and Marty Krofft, who created the H. R. Pufnstuf television show.[91] In 1970, Marty Krofft engaged in discussions and negotiations with the advertising agency Needham, Harper & Steers, Inc. about the possibility of basing a McDonald's advertising campaign on H. R. Pufnstuf.[92] Ultimately, Needham proceeded with the McDonald's advertising campaign and hired former employees of the Kroffts to do some of the work, but they did not license any rights concerning Pufnstuf.[93] When the campaign aired on television, the plaintiffs sued, claiming copyright infringement of the H. R. Pufnstuf episodes and various copyrighted articles of merchandise.[94] A jury returned a verdict for the plaintiffs, and McDonald's appealed.[95]

On appeal, the Ninth Circuit reformulated the test for copyright infringement. After citing *Reyher v. Children's Television Workshop* for the erroneously incomplete proposition that a copyright plaintiff must prove ownership of the copyright and copying by the defendant,[96] the court stated that copying is typically proven with access and substantial similarity.[97] This shows that the court began its analysis by repeating the mistake of applying improper appropriation as proof of copying.

---

[90] *See* Shyamkrishna Balganesh & Peter S. Menell, *The Use of Technical Experts in Software Copyright Cases: Rectifying the Ninth Circuit's "Nutty" Rule*, 35 BERKELEY TECH. L.J. 663, 676-77 (2020) (describing how *Krofft* "fundamentally misunderstood" *Arnstein*, in part by largely overlooking the issue of copying); Latman, *supra* note 51, at 1202-03 (noting "justifiable criticism" of *Krofft* and analyzing how *Krofft* "virtually assumes copying" and is "less helpful").

[91] *Sid & Marty Krofft Television Prods., Inc.*, 562 F.2d at 1160.

[92] *See id.* at 1161.

[93] *See id.*

[94] *See id.* at 1162.

[95] *See id.*

[96] *See id.* Unfortunately, this error even exists in Supreme Court jurisprudence. *See* Feist Publ'ns, Inc. v. Rural Tel. Serv., Inc., 499 U.S. 340, 361 (1991) (stating that infringement requires two elements of ownership and copying, omitting improper appropriation).

[97] *Sid & Marty Krofft Television Prods., Inc.*, 562 F.2d at 1162.

To its credit, the court criticized the use of substantial similarity as "boilerplate" that could lead to "untenable results," particularly overaggressive findings of infringement.[98] The court worried that, for example, the maker of a "cheaply manufactured plaster statue of a nude [person]" could successfully sue subsequent makers of similar objects because a jury might find the statues substantially similar (i.e., substantially similar in terms of their aesthetic appeal because they are both statues depicting nude persons).[99] The court tried to solve this problem by reformulating the test for copyright infringement. However, the supposed solution did not have the desired effect. Instead, it destroyed any sensible analysis of copying, making infringement depend entirely on improper appropriation.

According to the *Krofft* court, copyright infringement still required access and substantial similarity,[100] but substantial similarity now depended on a new two-part test. First, courts should conduct an "extrinsic" test to see if the defendant's work shares substantial similarity of ideas with the plaintiff's work.[101] Second, courts should conduct an "intrinsic" test to see if the two works share substantial similarity of expression.[102]

The extrinsic and intrinsic tests differed in their methods and objects of analysis. The court described extrinsic analysis as depending on "specific criteria that can be listed and analyzed."[103] These might include the type of artwork involved, materials used, subject matter, and the setting for the subject. The court clearly considered such similarities objective, often allowing decisions as a matter of law by a judge (presumably without factfinding by a jury).[104] By contrast, the intrinsic test depended on "the response of the ordinary reasonable person."[105] Here, specific formal similarities like material or setting no longer

[98] *Id.*

[99] *Id.*

[100] *Id.* (stating that infringement requires proof of ownership and "copying," and that "copying" is shown by access plus substantial similarity).

[101] *See id.* at 1164 (introducing extrinsic test).

[102] *See id.* (introducing extrinsic test).

[103] *Id.*

[104] *See id.*

[105] *Id.*

mattered.[106] Instead, observations and impressions of the ordinary observer took precedence.[107] The court asserted that this reformulation of copyright infringement pursued the same objectives as the *Arnstein* framework. It noted that under *Arnstein*, the trier of fact had to make determinations about copying and improper appropriation. The court then wrote:

> We believe that the court in Arnstein was alluding to the idea-expression dichotomy which we make explicit today. When the court in Arnstein refers to "copying" which is not itself an infringement, it must be suggesting copying merely of the work's idea, which is not protected by the copyright.[108]

With respect to the problem of overaggressive protection for things like statues, the Ninth Circuit apparently hoped to fashion a test that could separate actionable from non-actionable copying. Copyright denies protection to ideas,[109] so substantial similarity of ideas would seem non-infringing. By contrast, copyright protects the expression of ideas,[110] so substantial similarity of expression would seem infringing. Thus, an infringement claim by the court's hypothesized statue maker would fail because the statues in question would share only similarity of ideas.[111]

At first inspection, *Krofft*'s reformulation of infringement may seem effective and clever. Additional thought reveals, however, that *Krofft* effectively removed copying from the analysis of copyright infringement. *Krofft* therefore contradicted *Arnstein* and established

---

[106] *See id.* ("[The test] is intrinsic because it does not depend on the time of external criteria and analysis which marks the extrinsic test.").

[107] *Id.* (citing earlier case law for the proposition that the intrinsic test depends on the ordinary observer).

[108] *Id.* at 1165.

[109] 17 U.S.C. § 102(b).

[110] *See* Mazer v. Stein, 347 U.S. 201, 214 (1954) (stating that expressions, not ideas, are copyrightable); Baker v. Selden, 101 U.S. 99, 107 (1879) (holding that ideas in author's book are not protected by copyright).

[111] *Sid & Marty Krofft Television Prods., Inc.*, 562 F.2d at 1162-63 (identifying distinction between idea and expression to limit the possibility of claim over a "cheaply manufactured plaster statue of a nude").

copyright doctrine, even though the court apparently intended only to restate or clarify the law.

To illustrate this point, consider how a properly conducted *Arnstein* analysis would handle the statue hypothetical that troubled the *Krofft* court. Under *Arnstein*, if one sculptor sued another for copyright infringement simply because defendant's statue was also a "cheaply manufactured plaster statue of a nude [person]," the plaintiff would likely fail because the identified similarities could not support a finding of copying. Many have created such statues, making copying from the plaintiff a very poor explanation for the defendant's work. It is true that an alternate reason for finding against the plaintiff in this case would be improper appropriation — namely similarity only of ideas that do not rise to the level of infringement, but this determination would be made only *after* a finding of copying, and no such finding would need to be made.

In other words, the *Krofft* extrinsic/intrinsic framework eliminated the element of copying because substantial similarity of ideas does not establish copying. Two sculptures of nude figures may share substantial similarity of ideas (perhaps they both depict a man riding a horse-drawn chariot), but that does not establish that one sculptor copied from the other. The ideas might be common to many sculptures, derived from a third source like ancient Greek sculptures, or have been independently created. Similarity of ideas cannot establish copying unless the fact finder makes the further assessment that the ideas are unlikely to come from a common source or be independently created. Because *Krofft*'s extrinsic test does not require courts to make such a determination, it cannot be a test for copying, and it cannot be correctly equated to the first prong of the *Arnstein* framework.

As noted earlier, *Krofft* rivals *Arnstein* for its influence over the law of copyright infringement. Many courts have adopted *Krofft*'s extrinsic/intrinsic analysis. Predictably, their opinions perpetuate *Krofft*'s error and confusion.

For example, in the 1990 case of *Shaw v. Lindheim*,[112] the Ninth Circuit considered a claim by Shaw that the defendant's script for "The

[112] 919 F.2d 1353 (9th Cir. 1990).

Equalizer" infringed Shaw's script with the same title.[113] The district court granted the defendant's motion for summary judgment, and Shaw appealed.[114] Judge Alarcon's opinion began by stating that the plaintiff had to show copying, and that copying is shown by access plus substantial similarity.[115] Next, the opinion followed the *Krofft* framework, stating that substantial similarity depended on *Krofft*'s extrinsic/intrinsic analysis. This meant that the court could decide that the defendant copied without analyzing whether probative similarity existed. Unsurprisingly, the *Shaw* opinion suffers from this very problem.

The court noted that the district court granted summary judgment to the defendants on the ground that no reasonable juror could consider the two works substantially similar under a lay audience test.[116] The Ninth Circuit, however, believed that the subjective nature of the intrinsic lay audience test meant that judges could not use that test to grant summary judgment for a defendant.[117] Thus, a defendant could prevail at summary judgment only if the plaintiff failed to raise a triable issue of fact under the extrinsic test.[118] This led the court to analyze whether a reasonable juror could consider the two works substantially similar under the extrinsic test.[119]

To do this, the court listed a number of similarities between the two works identified by the plaintiff, including title, theme, twenty-six similar events, mood, setting, pace, characters, and dialogue.[120] The court considered these similarities sufficient to raise a triable issue of "substantial similarity" (in this context, substantial similarity of ideas).[121] This in turn meant a triable issue of infringement generally and, by implication, copying.

---

[113] *Id.* at 1355.

[114] *Id.*

[115] *Id.* at 1356.

[116] *Id.* at 1357-58.

[117] *Id.* at 1358 (citing *Krofft* for the proposition that the intrinsic test is "uniquely suited for determination by the *trier of fact*" (emphasis in original)).

[118] Shaw, 919 F.2d at 1359-60.

[119] *Id.* at 1361-64.

[120] *Id.*

[121] *Id.* at 1363-64.

The court's reasoning to support this conclusion is muddled and confused. Only once, in the consideration of similarities shared by the main characters, did the court conduct a recognizable analysis of probative similarity.[122] For the rest of the opinion, the court supported its analysis of copying with observations about aesthetic similarity, not probative similarity.

Consider how the court treated the fact that both scripts shared the theme of "a man who will equalize the odds, a lone man working outside the system."[123] The plaintiff wanted the court to infer that this similarity established copying, while the defendants argued that differences in their treatment of this theme justified a contrary inference.[124] A proper analysis of whether this similarity established copying would ask whether the common theme was more likely to exist by reason of copying, independent creation, or borrowing from a common source. This determination would in turn depend on the court's perception of how common or ordinary the shared feature was. However, the court did no such thing; it simply analyzed whether the themes of the two works were aesthetically similar. The court acknowledged the differences pointed out by the defendants, but these did not matter because they "did little to erode the similarity between the central themes embodied in the titles of the two works."[125] This amounts to nothing more than an assertion that the two works share substantial aesthetic similarity. The court never considered whether the similarities were sufficiently unlikely to support an inference of copying. Indeed, the court went on to quote a well-known adage that "[n]o plagiarist can excuse the wrong by showing how much of his work he did not pirate."[126] This adage may be true, but in the absence of any meaningful analysis of whether copying exists, it assumes that the defendant has already copied.

Similarly, when the court considered the twenty-six similarities of plot, the court recognized that these were largely "random similarities

[122] *Id.* at 1363 (discussing "literary accident" as an improbable explanation for similarities shared by the main characters).

[123] *Id.* at 1362.

[124] *Id.*

[125] *Id.*

[126] *Id.*

scattered throughout the works" and that Shaw "misrepresented" their order and similarity.[127] Under a proper analysis of copying, a court would consider whether these particular "random similarities" were sufficiently unusual to make their common occurrence probative of copying. However, the court failed to do this and simply asserted that the similarities existed. The court wrote, "Shaw's overexuberance, however, does not change the fact that many of the events in the two works are substantially similar."[128] This again meant that the court recognized a triable issue of copying without ever clearly making a probative similarity analysis.

To be clear, I am not claiming that the Ninth Circuit reached the wrong result. The plaintiff may well have deserved to survive summary judgment. Rather, I am pointing out that the court's reasoning fails to address the issue on which the decision should turn. Instead of identifying and assessing probative similarity as proof of copying, the court substituted aesthetic similarity for probative similarity. This is an error because, although aesthetic similarity is sometimes probative, it is not always so.[129] Accordingly, the opinion never explained why the plaintiff's identified similarities support a conclusion of copying. Thus, the court decided against the defendant without ever truly considering whether the defendant copied.[130]

C. The Current State of Copyright Infringement Law

The foregoing shows that although *Arnstein* provides a clear, analytically sound method for the circumstantial proof of copyright

[127] *Id.*

[128] *Id.* at 1363.

[129] *See supra* notes 51–60 and accompanying text.

[130] Other Ninth Circuit cases fail to distinguish probative from aesthetic substantial similarity. *See* Unicolors, Inc v. Urban Outfitters, Inc., 853 F.3d 980, 984-85 (9th Cir. 2017) (stating that copyright infringement may be proven with access and "substantial similarity" of idea and expression); Loomis v. Cornish, 836 F.3d 991, 994 (9th Cir. 2016) (holding that "absent direct evidence of copying, proof of infringement involves fact-based showings that the defendant had 'access' to the plaintiff's work and that the two works are 'substantially similar'"); Three Boys Music Corp. v. Bolton, 212 F.3d 477, 481 (9th Cir. 2000) (stating that a plaintiff may prove infringement by showing access and "substantial similarity"); *see also* Lemley, *supra* note 64, at 719-20 (criticizing the *Krofft* approach for, among other things, failing to pay proper attention to copying).

infringement, courts have not consistently implemented its framework. This has created caselaw littered with poorly reasoned opinions that easily lead judges to error. Although not every court errs, too many influential opinions lose sight of the important distinction between probative and aesthetic similarity.[131] Even when courts understand copying as a required element of copyright infringement, they often fail to analyze probative similarity, relying instead on aesthetic similarity. A line of Ninth Circuit music infringement opinions illustrates this phenomenon.

In the 2004 case of *Swirsky v. Carey*,[132] the court considered a claim that the defendant Mariah Carey's "Thank God I Found You" infringed Seth Swirsky and Warryn Campell's "One of Those Love Songs."[133] The district court granted summary judgment to Carey, and the plaintiffs appealed.[134] In reversing, the Ninth Circuit adopted the *Krofft* framework.[135] Doing so meant that the court could find evidence sufficient to withstand summary judgment simply by identifying aesthetic similarities between the two works without analyzing whether the similarities were probative of copying. And indeed, this is precisely what happened. Critical to this was the testimony of the plaintiffs' expert, who testified that "the two songs had substantially similar choruses."[136] As the court described it, the expert's methodology identified "how the two choruses sound[ed] to his expert ears."[137] This

[131] *See* Batiste v. Najm, 28 F. Supp. 3d 595, 599 (E.D. La. 2014) (noting confusion about use of term "substantial similarity" and distinguishing probative similarity from aesthetic similarity); *see also* Positive Black Talk Inc. v. Cash Money Records Inc. 394 F.3d 357, 370 (5th Cir. 2004) ("In order to avoid confusion, a district court should explain that the purpose of the probative similarity inquiry is to determine whether factual copying may be inferred and that this inquiry is not the same as the question of substantial similarity, which dictates whether the factual copying, once established, is legally actionable.").

[132] 376 F.3d 841 (9th Cir. 2004).

[133] *Id.* at 843-44.

[134] *Id.* at 844.

[135] *Id.* at 845 (using extrinsic and intrinsic tests to analyze substantial similarity).

[136] *Id.* at 845.

[137] *Id.* at 847.

is evidence of aesthetic similarity, not probative similarity, but the court nevertheless relied on it to remand the case for trial.[138]

By adopting the *Krofft* framework, *Swirsky* necessarily replicated *Krofft*'s analytical errors. However, the Ninth Circuit appeared to recognize the error of its ways in the 2020 case of *Skidmore as Trustee for Randy Craig Wolfe Trust v. Led Zeppelin*.[139] In *Skidmore*, the plaintiff claimed that the Led Zeppelin classic "Stairway to Heaven" infringed Randy Wolfe's "Taurus."[140] Skidmore claimed that the opening notes of "Stairway to Heaven," particularly the famous A minor chord progression, infringed the opening of "Taurus."[141] After a full trial, the jury returned a verdict for Led Zeppelin, and Skidmore appealed.[142] The Ninth Circuit, sitting en banc, affirmed.[143]

In so doing, the Ninth Circuit summarized the law of copyright infringement. It clearly identified "copying" and "unlawful appropriation" as requirements, specifically noting the distinction between probative and aesthetic similarity.[144] This led to the declaration that both the extrinsic and intrinsic tests should be used to establish unlawful appropriation, not copying.[145] Copying had to be proven, but with probative similarity, not aesthetic substantial similarity.[146]

The *Skidmore* opinion clearly understood that courts should not analyze infringement the way *Krofft* and *Swirsky* did. *Skidmore* could therefore represent a salutary effort to fix the problems created when courts lose sight of the distinction between probative and aesthetic similarity. However, a mere two years later, the Ninth Circuit once again lost its way.

---

[138] *Id.* at 853 (plaintiffs' expert provided sufficient evidence of disagreement about the substantial similarity of the works to survive summary judgment).

[139] 952 F.3d 1051 (9th Cir. 2020).

[140] *Id.* at 1056.

[141] *Id.* at 1057-58.

[142] *Id.* at 1060.

[143] *Id.* at 1079.

[144] *Id.* at 1064 (the term "substantial similarity" embodies distinct concepts directed towards copying and unlawful appropriation).

[145] *Id.* ("unlawful appropriation" depends on substantial similarities, which are proven by use of the extrinsic and intrinsic tests).

[146] *Id.* (copying is proven circumstantial through access and probative similarity).

In *Gray v. Hudson*,[147] the court ruled on a claim in which Marcus Gray alleged that Katy Perry's song "Dark Horse" infringed his song "Joyful Noise."[148] Here, the district court vacated a jury award and granted judgment as a matter of law to the defendants.[149] The Ninth Circuit affirmed and again summarized the law of infringement. The court began by stating that the plaintiff had to show copying, and that a circumstantial case of copying required access and substantial similarity.[150] It then cited *Skidmore* for the proposition that substantial similarity, and in the context of this case, copying, would be determined by applying the extrinsic and intrinsic tests.[151]

Despite the citation to *Skidmore*, one cannot read *Gray*'s summary of infringement law as consistent with *Skidmore*'s. *Skidmore* carefully distinguished probative from aesthetic similarity, applying probative similarity to the question of copying and aesthetic similarity to the question of improper appropriation.[152] By contrast, *Gray* repeated the error of forgetting this distinction and collapsed the two forms of similarity in a way that again made the analysis of probative similarity unnecessary.[153] Indeed, other courts have already cited *Gray* for the proposition that a plaintiff may prove copying with access plus substantial similarity, and that the extrinsic and intrinsic tests determine whether substantial similarity exists.[154]

Not surprisingly, this pattern of case law exists beyond the Ninth Circuit,[155] and it creates a complicated backdrop against which to study

[147] 28 F.4th 87 (9th Cir. 2022).

[148] *Id.* at 92.

[149] *Id.*

[150] *Id.* at 96.

[151] *Id.*

[152] *See supra* notes 139–146 and accompanying text.

[153] *See supra* notes 147–151 and accompanying text.

[154] *See, e.g.*, Kahn v. CJ E & M America, Inc., No. CV 21-3230-DMG (KSx), 2022 WL 2037495, at *4 (C.D. Cal. Mar. 28, 2022).

[155] *See* Ellis v. Diffie, 177 F.3d 503, 506 (6th Cir. 1999) (recognizing copying as an element of infringement that depends on "substantial similarity" and referring to the "ordinary observer" test as the "traditional standard" for substantial similarity); Moore v. Columbia Pictures Industries, Inc., 972 F.2d 939, 941-45 (8th Cir. 1992) (copying may be established by access and "substantial similarity" before citing *Krofft* for proposition that substantial similarity involves similarity of idea and expression); Atari Games Corp.

the admissibility of forensic musicology opinions. As the next Part will discuss, courts have long been aware that expert testimony can cause problems in copyright infringement litigation. However, confusion about the distinction between probative and aesthetic similarity has opened the door to questionable uses of forensic musicology.

II. FORENSIC MUSICOLOGY TESTIMONY AND THE FEDERAL RULES OF EVIDENCE

A. The Content of Forensic Musicology Opinions

Applying the Federal Rules of Evidence to forensic musicology testimony begins by identifying four kinds of testimony commonly found in a forensic musicology opinion. First, the expert identifies formal musical features shared by the works in question.[156] Such testimony draws the jury's attention to formal similarities that the parties consider important. Second, experts offer information about identified similarities to affect their evidentiary impact.[157] For example, describing a shared musical feature as rare or a standard compositional characteristic may affect whether a jury thinks that the defendant copied from the plaintiff or the plausibility of the plaintiff's claim of ownership.[158] Similarly, testimony about the musical significance of a

v. Nintendo of America, Inc., 975 F.2d 832, 837-38 (Fed. Cir. 1992) (referring to copying of ideas and then expression).

[156] *See infra* notes 162–167 and accompanying text.

[157] *See infra* notes 168–172 and accompanying text.

[158] *See* 17 U.S.C. § 102(a) (limiting copyright protection to "original works of authorship"). The common existence of a musical feature in works predating the plaintiff's work raises the possibility that the plaintiff actually copied the relevant feature from someone else. If the jury believed this, then the jury could not find the feature original to the plaintiff. However, it is important to understand that the pre-existence of a musical feature does not establish that future works containing this feature necessarily arose from copying. Coincidence remains an entirely plausible explanation. However, that the common existence of a musical feature surely weakens the argument that the defendant must have copied that feature from the plaintiff's work. First, common existence provides exemplars of other composers who have written works with the same feature, thereby strengthening the defendant's argument that coincidence explains the similarities. Second, common existence raises the possibility that if the defendant copied, he did so by copying from a work other than the plaintiff's.

shared feature may influence the jury's appraisal of aesthetic similarity and improper appropriation. Third, experts characterize the degree of similarity demonstrated by the identified musical features. This ranges from the evaluation of a few shared notes to a cumulative opinion about the overall similarity of entire works, and it often includes a conclusion that the two works are or are not "substantially similar."[159] Fourth, experts often opine that the identified similarities prove that the defendant did (or did not) copy from the plaintiff's work.[160] Although the general content of such testimony is easy to imagine, it is worth providing examples of expert testimony given in recent headline-making trials so that later evaluation against the Federal Rules of Evidence will be clear.[161]

Testimony offered by the plaintiff in *Gray v. Perry* demonstrates how an expert identifies musical similarity. The expert defined "ostinato" as "a repeating figure that cycles within a piece."[162] He opined that the works in question shared a similar ostinato.[163] He then defined a "scale

Because of this, the Article will treat evidence of common occurrence as primarily relevant to the question of copying, as opposed to ownership.

[159] *See infra* note 173 and accompanying text.

[160] *See infra* note 174 and accompanying text.

[161] The three cases are: Gray v. Perry, No. 2:15-CV-05642-CAS-JCx, 2020 WL 1275221 (C.D. Cal. Mar. 16, 2020) (claim of infringement against pop star Katy Perry concerning "Dark Horse"); Skidmore v. Led Zeppelin, No. CV 15-03462-RGK (AGRx), 2016 WL 6674985 (C.D. Cal. Aug. 8, 2016) (claim of infringement against famous rock band Led Zeppelin concerning "Stairway to Heaven"); Williams v. Bridgeport, No. LA CV13-06004-JAK (AGRx), 2014 WL 7877773 (C.D. Cal. Oct. 30, 2014) (claim of infringement against Pharell Williams and Robin Thicke concerning "Blurred Lines").

[162] Testimony of Plaintiff's Expert Witness, Todd Ryan Decker, Ph.D., Gray v. Perry, No. 2:15-CV-05642r-CAS-JCx, 2019 WL 10982151 (C.D. Cal. July 19, 2019).

> Q. What is an ostinato?
>
> A. An ostinato is a repeating figure that cycles within a piece. Um, in this case it's eight notes long. Uh, it's a repeating figure that cycles and upon which the structure of the piece is built. So one analogy to use would be, um, sections of toy train track. The ostinato is a section and you line up sections of that train track and build the piece upon those repetitions of the ostinato.

[163] *Id.*

> Q. I'd like to now shift to talk about the details of your opinion. In your report you said the most obvious or pervasive similarity is the ostinato; is that right?

degree" as "a way to assign a number to a pitch in a scale."[164] This allowed the expert to use numbers while singing the two ostinatos in order to demonstrate to the jury that, although the ostinatos did not use identical notes (because they were written in different keys), they shared what the expert considered many musically identical pitches.[165]

A. Yes.

[164] *Id.*

Q. Well, wait. First, can you tell us what is a scale degree?

A. So a scale degree is a number that identifies where a given pitch — it's a way to assign a number to a pitch in a scale. So one would be the tonic or the home note. Five would be the dominance for those with some musical training and so forth.

So there are seven pitches in a scale and so it would seven scale degrees. Scale degrees are indicated in music theory with a number and a little caret above it so that's why I use that caret. That's music theory notations or lingo.

[165] *Id.* The testimony explaining the use of scale degrees was as follows:

Q. Thank you. And what is the purpose of using scale degrees?

A. So scale degrees allow you to see through the differences in letter names. For example, if you sing the song Happy Birthday and you sing it in C major, it would go C C D C F G. If you sing it in G major, it would go G G A G C B. Sounds like two completely different melodies in that respect, but we all know that those are in fact the same melody starting at different starting point.

So a way to sing Happy Birthday that wipes away those — that elements letter names is to sing degrees. 1 1 2 1 4 3. You can start on any pitch 1 1 2 1 4 3, 1 1 2 1 4 3, anywhere, and you have a sense for the relationship of the notes in that melody. The relationship of the pitches . . . .

The expert's use of scale degrees to support his opinion about the existence of musically identical pitches is as follows, with ellipses indicating omission of questions and testimony clarifying the initial question:

Q. And did you make any determination regarding, um, the similarity of the pitch?

A. [T]his melody starts at 3, the third scale degree and over the course and remains on that third scale degree for four notes, four iterations, four playings of the note. Then it steps down to 2, and then it goes to 1. So both of the melodies have this 3, 2, 1 gesture.

And then the 6 and 5 at the end have to do with a leap down. So while those notes are higher, our ear they're in a lower octave so they are actually a jump

Similarly, in *Skidmore v. Led Zeppelin*, the defendant's expert described a chord progression at issue in the litigation as a "chromatic descending minor chord progression."[166] The expert then identified the leading notes for the progressions, starting with A and moving to F.[167]

These cases also illustrate how experts offer information about similarities. In *Gray v. Perry*, the plaintiff's expert stated that he had not seen another work with ostinatos like the one he had identified.[168] He further characterized the even rhythm of the ostinatos as "unusual" because American popular music typically features syncopation.[169] And

> down. So to sing it again, 3 3 3 3 2 2 2 1, 3 3 3 3 2 2 2 6. There's a leap there that then we jump back up to 3 to restart the ostinato.
>
> This also happens in Dark Horse. 3 3 3 3 2 2 1 5, 3 3 3 3 2 2 1 5. So that, that descending melodic trajectory and also the fact the melody sits on 3 for four playings is — is noteworthy.

[166] Reporter's Transcript of Jury Trial, Day 4, Vol. 1 at 809, Skidmore v. Led Zeppelin, No. CV 15-03462-RGK, 2016 WL 6674985 (C.D. Cal. June 17, 2016).

> Q. What are the similarities between the Section A of "Taurus" that — as you've described it, and "Stairway to Heaven"?
>
> A. The similarities are limited. It's a similarity, in fact. It's a similarity of a common chromatic descending minor chord progression. And I think we drop the "progression" because it's so long already, so I think just simply call it a descending chromatic minor line progression.

[167] *Id.* at 810.

[168] Testimony of Plaintiff's Expert Witness, Todd Ryan Decker, Ph.D., *supra* note 162.

> Q. Do you see this often with an eight note ostinato or beat such as this that you would see commonly that a beat descends in this fashion?
>
> A. I have not seen another, a third piece that descends in the way these two do.

[169] *Id.*

> Q. So as far as the phrase length goes, I think you already testified that there are eight beats in each —
>
> A. Yes.
>
> Q. — in each ostinato. Um, is that characteristic in contemporary beat driven music?

in *Skidmore v. Led Zeppelin*, the defendant's expert characterized the descending chord progression as "commonplace" and a "musical building block."[170] With respect to aesthetic significance, in *Williams v. Bridgeport*, the plaintiff's expert identified the "signature phrase" of a work as important[171] and a song's "hook" as "the most well-known or memorable theme of a song."[172]

Finally, to see how experts go from providing basic information to offering conclusions about similarities, consider this exchange from *Gray v. Perry*:

> Q. And so we're going to get into the details of your opinion, but to summarize based on your expertise and music performance, musical history and musical borrowing, do you think Joyful Noise and Dark Horse are substantially similar?

> A. Sure. It's characteristic for a phrase like this to last for eight beats. What's distinctive here is that rhythm of eight beats is completely even. All eight of those beats for both ostinatos are completely the same.
>
> So a way to demonstrate that would be to clap this, um, clap the ostinato. So I'll sing both ostinatos and clap. 3 3 3 3 2 2 1 6, 3 3 3 3 2 2 1 5. So if I just clap the rhythm of these ostinatos without singing.
>
> (Witness clapping.)
>
> THE WITNESS: Completely even, all of the values are the same, and they unfold, um, with no syncopation. For example, these are squarely on the beat which is unusual. In popular music in America for the last 150 years, syncopation has been the thing that says American music and that's lacking here.
>
> BY MS. COHEN:
>
> Q. What is syncopation?
>
> A. Syncopation is a melody that has notes or accents that are off the beats.

[170] Reporter's Transcript of Jury Trial, Day 4, Vol. 1, *supra* note 166, at 811 (characterizing similarity as "commonplace" and a "musical building block").

[171] Testimony of the Plaintiff's Expert Witness, Williams v. Bridgeport, No. CV-13-06004-JAK (AGRx), 2015 WL 4742406 (C.D. Cal. Mar. 12, 2015) (expert opining that certain similarities occur in the "signature phrase" which is "identifying or very recognizable"). The transcript in question is entitled "Testimony of the Defendant's Expert Witness," but this appears to be an error.

[172] *Id.*

A. I do.[173]

Note also how the expert followed this aesthetic appraisal with a cumulative assertion about copying:

Q. Do you think Dark Horse borrows from Joyful Noise?

A. I do.[174]

B. Basic Principles of Evidence Law

A brief exposition of basic evidence law will allow us to evaluate the uses of forensic musicology described above. The Federal Rules of Evidence govern the admissibility of evidence in all United States District Courts and Courts of Appeals.[175] The Rules exist to "promote the development of evidence law, to the end of ascertaining the truth and securing a just determination."[176] Accordingly, the Rules admit evidence that supports reasoned deliberation by juries while excluding evidence that distracts, misleads, or confuses.

The foundational principles behind the Rules are relevance and prejudice, embodied respectively in Rules 401 and 403. Pursuant to Federal Rule of Evidence 402, only relevant evidence may be admitted at trial. Rule 401 defines evidence as relevant if:

---

[173] Testimony of Plaintiff's Expert Witness, Todd Ryan Decker, Ph.D., *supra* note 162; *see also* Testimony of the Plaintiff's Expert Witness, *supra* note 171 (plaintiff's expert asserting that two works are "significantly similar").

[174] Testimony of Plaintiff's Expert Witness, Todd Ryan Decker, Ph.D., *supra* note 162; *see also* Trial Transcript at 304, Griffin v. Sheeran, No. 17-CV 5221 (LLS) (plaintiff's expert opining that defendants copied from the plaintiff's work); Testimony of the Plaintiff's Expert Witness, Williams v. Bridgeport, *supra* note 171 (expressing plaintiff's expert's opinion that the works expressed "shared creative choices" that were not the result of coincidence); Trial Transcript, *supra* note 174, at 655-56, (expressing defendant's expert's "very strong opinion" that defendants did not copy because "the chord progressions were in common use . . . and as a result, there's no musicological evidence of copying"); Repp v. Webber, 132 F.3d 882, 886-87 (2d Cir. 1997) (referring to opinion of expert witness that similarities between works were so great "that I cannot consider them as insignificant or coincidental, and I must conclude that [one is based on the other]").

[175] FED. R. EVID. 1101(a).

[176] FED. R. EVID. 102.

> (a) it has any tendency to make a fact more or less probable than it would be without the evidence; and
>
> (b) the fact is of consequence in determining the action.[177]

Importantly, however, relevance alone does not guarantee admissibility. When otherwise relevant evidence presents an undue risk of unfair prejudice, the Rules give courts discretion to exclude such evidence. Under Rule 403, courts properly exercise such discretion when the probative value of evidence is "substantially outweighed" by the danger of prejudice, confusion, or misleading the jury.[178]

The Advisory Committee Notes to Rule 403 elaborate that evidence is unfairly prejudicial when it has "an undue tendency to suggest decision on an improper basis."[179] Thus, for example, in *Old Chief v. U.S.*,[180] the Supreme Court reversed a lower court ruling which failed to properly determine whether the probative value of marginally relevant evidence was substantially outweighed by its danger of prejudice. In that case, the defendant Old Chief appealed his conviction for unlawful possession of a firearm by a person with a prior felony conviction.[181] At trial, the prosecution proved, over Old Chief's objection, that Old Chief had been convicted of a felony assault causing serious bodily injury.[182] Old Chief objected because he did not want the jury to know the particulars of his earlier conviction, and he offered to stipulate to having been convicted of a felony. If accepted, the jury would learn that Old Chief was a convicted felon, but not that he had been convicted of a violent crime.[183] The government refused to stipulate, and the trial court allowed the jury to hear about Old Chief's conviction for assault.[184] Old Chief appealed, arguing that the specific nature of his felony conviction was unduly

[177] FED. R. EVID. 401.

[178] FED. R. EVID. 403.

[179] FED. R. EVID. 403 advisory committee's note.

[180] 519 U.S. 172 (1997).

[181] *See id.* at 174-77.

[182] *See id.* at 175-77.

[183] *See id.* at 175-76.

[184] *Id.* at 177.

prejudicial, but the Ninth Circuit affirmed.[185] The Supreme Court granted certiorari and reversed.[186]

In so deciding, the Court noted that Old Chief's conviction for assault (as opposed to simple conviction of a felony) was relevant because it proved an element of the crime.[187] However, the evidence was also unduly prejudicial because it invited the jury to conclude that the defendant had a violent character that made him likely to commit other crimes.[188] The jury might then convict the defendant because of his perceived character or propensity, as opposed to specific evidentiary proof. The Court explained that "'unfair prejudice' speaks to the capacity of some concededly relevant evidence to lure the factfinder into declaring guilt on a ground different from proof specific to the offense charged."[189] Thus, the trial court should have excluded the evidence because it raised an unacceptable risk of leading the jury to decide the case on an improper basis (namely the supposed bad character of the defendant).[190]

The Rule 401/403 balance also applies to expert testimony. For example, in *United States v. Fallon*,[191] three defendants appealed convictions for criminal fraud.[192] Among the issues on appeal was the district court's exclusion of expert testimony by a law professor offered for the purpose of demonstrating that certain contracts and documents could be interpreted in a manner consistent with the defendant's good faith.[193] The Third Circuit affirmed under Rule 403.

At trial, the district court raised concerns that the professor's testimony could "turn this case into a contract case even though it's a

[185] *See id.* at 177.

[186] *Id.* at 177-78.

[187] *See id.* at 178-79.

[188] *See id.* at 180-81.

[189] *Id.* at 180.

[190] *Id.* at 181; *see also* United States v. Delgado-Marrero, 744 F.3d 167 (1st Cir. 2014) (finding evidence of sexual orientation improper under Rule 403) Candonoza v. Royston, No. 4:16-CV-0048-HLM, 2018 WL 5023425, at *3 (D. Ga. 2018) (finding sexual orientation properly excluded under Rule 403 as unduly prejudicial).

[191] 61 F.4th 95 (3d Cir. 2023).

[192] *Id.* at 103.

[193] *Id.* at 108.

fraud case."[194] The Third Circuit agreed, finding that a "well-credentialed law professor" would lead the jury to believe that the case turned on interpretation of the contracts instead of the defendants' subjective intent.[195] Thus, although expert interpretation of the contracts could provide circumstantial evidence about what the defendants might have believed, this probative value was "substantially outweighed by the risk of confusing the jury on the role that contract law played in the dispute."[196]

In addition to Rules 401 and 403, which apply to all forms of evidence, the Federal Rules of Evidence also contain specific provisions concerning expert testimony.[197] These provisions exist because expert testimony poses peculiar risks to the fact-finding process.[198] Expert witnesses testify because some cases raise issues that lay jurors do not understand.[199] Jurors rely on expert testimony to understand these cases and shape deliberation, but they also lack the knowledge or experience to evaluate whether such testimony is well-founded.[200] Mistaken, unreliable, or exaggerated expert testimony is therefore more likely to mislead a jury than ordinary testimony about subjects that jurors already understand well.[201]

---

[194] *Id.* at 109.

[195] *Id.* at 109-10.

[196] *Id.* at 110.

[197] FED. R. EVID. 702-06.

[198] *See* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 595 (1993) (noting that trial judges should exercise more control over expert testimony than lay testimony because jury may have difficulty evaluating expert testimony); Edward J. Imwinkelried, *The Epistemological Trend in the Evolution of the Law of Expert Testimony: A Scrutiny at Once Broader, Narrower, and Deeper*, 47 GA. L. REV. 863, 871-72 (2013) (discussing risk that expert testimony, especially if based in science, may "overawe lay jurors and mislead them into ascribing undue weight to the testimony").

[199] *See* Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149 (1999) (recognizing that an expert's testimony often rests upon knowledge that juries do not have).

[200] *Id.* (ensuring reliability of expert testimony helps jury evaluate that testimony).

[201] *See Daubert*, 509 U.S. at 595 (noting that expert testimony can be "both powerful and quite misleading"); Clark v. Edison, 881 F. Supp. 2d 192, 215 (D. Ma. 2012) (referring to risk that expert testimony admitted under Federal Rule of Evidence 702 may unduly influence the jury); United States v. Mejia, 545 F.3d 179, 190-91 (2d Cir. 2008) (describing risk that an "expert transforms into the hub of the case, displacing the jury . . .").

The Federal Rules of Evidence manage this problem primarily through Rule 702,[202] which provides as follows:

> A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:
>
> (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
>
> (b) the testimony is based on sufficient facts or data;
>
> (c) the testimony is the product of reliable principles and methods; and
>
> (d) the expert has reliably applied the principles and methods to the facts of the case.[203]

Rule 702 tries to prevent expert testimony from going awry through three different requirements. First, the expert must be properly qualified, thereby ensuring (at least in theory) a knowledgeable witness. Second, under subsection (a), the proposed testimony must be helpful to the jury. This excludes expert testimony on subjects that jurors adequately comprehend from a lay perspective.[204] Third, under subsections (b)–(d), the proposed expert testimony must be reliable. This implies rejecting expert testimony based on personal hunch or

---

[202] *See, e.g.*, McCorvey v. Baxter Healthcare Corp., 298 F.3d 1253, 1256 (11th Cir. 2002) (stating that courts act under Rule 702 "to ensure that speculative unreliable expert testimony does not reach the jury").

[203] FED. R. EVID. 702.

[204] *See* United States v. Lespier, 725 F.3d 437, 449 (4th Cir. 2013) (Federal Rule of Evidence 702 prohibits expert testimony obviously within the common knowledge of jurors); United States v. Frazier, 387 F.3d 1244, 1262 (11th Cir. 2004) (stating that, under Federal Rule of Evidence 702, "expert testimony is admissible if it concerns matters that are beyond the understanding of the average lay person"); Stromback v. New Line Cinema, 384 F.3d 283, 295 (6th Cir. 2004) (noting that expert testimony will seldom be necessary in copyright infringement cases about works aimed at a general audience); 4 WEINSTEIN'S FEDERAL EVIDENCE § 702.03 (2024) ("Expert testimony is generally not permitted concerning factual issues that are within the knowledge and experience of ordinary lay people.").

opinion, as opposed to reasoned conclusions based on objective data, established principles, and tested methods of inquiry.[205]

The basic requirement of expert qualification makes intuitive sense. Given the risk of misleading juries, it is important to ensure that expert witnesses know enough about their subject area to meaningfully help juries. So far, there have not been serious issues about the qualification of forensic musicology witnesses, making it unnecessary to further elaborate on this requirement.

The requirement of helpfulness exists to ensure that experts do not upset or interfere with deliberations that the jury is adequately qualified to conduct on its own. Because experts often opine about issues that a jury will evaluate, there is a risk that the jury will rely on the expert's judgment instead of deliberating independently.[206] This may be inevitable when the expert testifies about things the jury does not understand well enough to evaluate, but it is avoidable when expert testimony concerns things the jury understands well enough to deliberate over. Rule 702 therefore excludes expert testimony about matters that juries can adequately comprehend on their own.[207]

For example, consider *Sonos, Inc. v. D & M Holdings*,[208] a patent infringement case in which both sides retained multiple experts.[209] Prior to trial, the defendant D & M moved to strike portions of a Sonos expert's report stating that D & M had copied Sonos' patents.[210] The trial court granted the motion on the basis of Rule 702(a) because jurors were fully capable of evaluating the relevant evidence of copying on their own.[211] The opinion of copying would therefore not help the jury.

---

[205] *See Daubert*, 509 U.S. at 590 (stating that requirement of knowledge in Rule 702 "connotes more than subjective belief or unsupported speculation").

[206] *See supra* notes 198–201 and accompanying text.

[207] *See* FED. R. EVID. 702 advisory committee's notes (noting that expert testimony is not admissible when the "untrained layman" can appropriately understand the issue); United States v. Baca, 2018 U.S. Dist. LEXIS 211943, at *56 (D. N.M. 2018) ("[E]xpert testimony is proper only if a typical juror would be unable to fully understand without expert assistance.").

[208] 297 F. Supp. 3d 501 (D. Del. 2017).

[209] *See id.* at 506 (identifying experts for each party).

[210] *See id.* at 520-21.

[211] *See id.* at 520-22 (citing Rule 702 and concluding that jury was fully capable of understanding evidence of copying without expert's assistance).

Instead, it would function as argument that the jury might rely upon because it carried the "special imprimatur" of presentation by an expert.[212]

The final and most important Federal Rules of Evidence limit on expert testimony is reliability. As the Advisory Committee notes make clear, Rule 702 codifies two foundational Supreme Court precedents, *Daubert v. Merrell Dow Pharmaceuticals*[213] and *Kumho Tire Co., Ltd v. Carmichael*.[214]

In *Daubert*, the infant plaintiffs and their guardians ad litem sued Merrell Dow to recover for birth defects allegedly caused by their mothers' use of Bendectin to treat nausea.[215] After discovery, Merrell Dow moved for summary judgment on the ground that no published study had found Bendectin to cause birth defects.[216] The plaintiffs conceded that this was true. Nevertheless, they argued against summary judgment on the basis of eight expert witnesses who had concluded that Bendectin does cause birth defects.[217] These experts held their belief on the basis of animal and *in vitro* studies finding a link between Bendectin and birth defects, as well as chemical similarities between Bendectin and other known teratogens.[218] The district court granted Merrell Dow's motion for summary judgment, rejecting the plaintiffs' expert opinions as inadmissible because their conclusions were not generally accepted by the relevant scientific community.[219] The Ninth Circuit affirmed.[220]

The Supreme Court granted certiorari, reversed, and remanded.[221] In so ruling, the Court found that the then-prevalent standard of general acceptance used by the trial court had been superseded by the Federal

[212] *Id.* at 522 (stating that Sonos may present evidence of copying to the jury, but not "by invoking the special imprimatur" that accompanies expert testimony).

[213] 509 U.S. 579 (1993).

[214] 526 U.S. 137 (1999); FED. R. EVID. 702 Advisory Committee's Notes (noting 2000 amendment of Rule 702 in response to *Daubert* and *Kumho Tire*).

[215] *Daubert*, 509 U.S. at 582.

[216] *Id.*

[217] *Id.* at 583.

[218] *Id.*

[219] *Id.* at 584.

[220] *Id.*

[221] *Id.* at 597-98 (vacating judgment of Court of Appeals and remanding).

Rules of Evidence.[222] Because Federal Rule of Evidence 702 made no mention of general acceptance, the Court held use of the general acceptance standard was inappropriate.[223] Instead, courts considering expert testimony based on science should embark upon a flexible inquiry designed to ensure the reliability and relevance of any science backing the expert's opinion.[224] A trial judge should first assess whether the reasoning or methodology supporting the opinion is valid and whether it can be applied to the case at hand.[225] This ordinarily includes consideration of whether the method or principles have been tested, the existence of peer review, known or potential rates of error, and general acceptance.[226] A judge should further consider whether the expert testimony has undue potential to mislead the jury and whether, under Rule 403, this prejudice substantially outweighs any probative value.[227] The Court then remanded the case for further consideration.[228]

The Court elaborated its interpretation of Rule 702 in *Kumho Tire Co., Ltd. v. Carmichael*,[229] extending *Daubert* to all expert testimony, not just scientifically based expert testimony.[230] In *Kumho Tire*, the Carmichaels brought a diversity action against Kumho Tire for damages associated with a car accident that happened when a tire on the minivan driven by Patrick Carmichael blew out.[231] The plaintiffs asserted that the tire blew out because it was defective, and they sought to prove defectiveness with the testimony of an expert in tire failure analysis, Dennis Carlson, Jr.[232]

Carlson opined that a defect caused the blowout in part by ruling out the possibility that underinflation or overloading had caused

[222] *Id.* at 587.

[223] *Id.* at 588.

[224] *Id.* at 597 (Federal Rules of Evidence "assign to the trial judge the task of ensuring that an expert's testimony both rests on a reliable foundation and is relevant to the task at hand").

[225] *Id.* at 592.

[226] *Id.* at 592-94.

[227] FED. R. EVID. 403; *Daubert*, 509 U.S. at 595.

[228] *Daubert*, 509 U.S. at 598.

[229] 526 U.S. 137 (1999).

[230] *Id.* at 141-42.

[231] *Id.* at 142.

[232] *Id.*

"overdeflection" that separated the tire tread from the rest of the tire.[233] Carlson ruled out this possibility because, according to him, a tire subject to overdeflection should exhibit certain physical signs that include a certain pattern of tread wear, signs of a "bead groove" caused by pushing the tire bead too hard against the inside of the tire's rim, physical signs of deterioration on the tire sidewalls, and marks on the tire's rim flange.[234] Carlson testified at his deposition that a defect is the probable cause of a tire blowout unless he finds at least two of the physical signs upon the relevant tire.[235] Carlson conceded that signs of suspicious tread wear, bead groove, and marks on the rim flange existed.[236] However, he believed that these signs were insignificant, and he therefore discounted them as evidence of overdeflection.[237]

Kumho Tire moved to exclude Carlson's testimony as not sufficiently reliable under Rule 702.[238] The district court responded by applying *Daubert* to the case at hand and concluded that Carlson's testimony was indeed unreliable, in part because his methodology could not be tested for validity, had not been subject to peer review or publication, had no sufficient known rate of error, and lacked acceptance within the relevant community.[239] The Eleventh Circuit reversed, stating that *Daubert* covered only scientific expert testimony, and that Carlson's testimony was admissible as relying on experience, as opposed to science.[240] Kumho Tire petitioned for certiorari, asking the Supreme Court to determine if a trial court "may" consider *Daubert*'s factors in non-science cases.[241]

The Court found that Federal Rule of Evidence 702 applied to all forms of expert testimony, and that *Daubert*'s general principles fully applied to the case at hand.[242] Thus, trial courts may (not must) consider

[233] *Id.* at 143-45.
[234] *Id.* at 144.
[235] *Id.*
[236] *Id.* at 144-45.
[237] *Id.*
[238] *Id.* at 145.
[239] *Id.*
[240] *Id.* at 146.
[241] *Id.*
[242] *Id.* at 147-49.

factors such are testing, peer review and publication, known error rates, standards controlling an expert's determinations, or general acceptance.[243] This meant that for all expert testimony, trial courts must ensure reliability and relevance even if they have flexibility in deciding how to go about the task at hand.[244] The Court then found the district court's exclusion of Carlson's testimony reasonable and reversed the Eleventh Circuit.[245]

For our purposes, *Kumho Tire* is particularly important because it illustrates the requirement of reliability for expert opinions not based solely on science. Expert testimony does not become admissible simply because the proffered opinion is plausible. The testimony must also rest upon appropriately validated methodology or experience.[246] This ensures that the expert's testimony is reliable and not a matter of personal opinion.

Carlson's assertion about the tire blowout was plausible. One can easily imagine how overdeflection could leave signs of wear that an experienced engineer could spot. Indeed, Carlson's conclusion about defect as the cause might have been correct. Nevertheless, the Supreme Court held that trial court properly excluded Carlson's testimony.[247] Carlson's opinion may have been plausible, but it was not reliable because it depended too heavily on Carlson's personal fiat.

As the Court explained, Carlson's methodology had no basis in engineering. [248] He did not establish that the factors he considered were those that other experts would analyze, nor did he explain why two factors (as opposed to one or three) established overdeflection.[249] There was no testing, peer review, controlling standard of uniformity, or general acceptance.[250] Even if one were to accept Carlson's general

[243] *Id.* at 149-50.

[244] *Id.* at 150.

[245] *Id.* at 158.

[246] *Id.* at 149 (requiring expert testimony to have a reliable basis in the knowledge and experience of a relevant discipline).

[247] *Id.* at 158.

[248] *Id.* at 157 (noting that absence of evidence that other experts would use Carlson's methods for detecting and analyzing possible tire defects as well as the absence of studies validating his approach).

[249] *Id.*

[250] *See id.*

methodology, Carlson could not explain why he was correct to dismiss evidence of overdeflection as insubstantial other than by relying on his personal judgment.[251] Carlson might assert that his methods and conclusions were accurate, but that did not establish sufficient reliability because the claim stood primarily on his personal say-so.[252]

C. Forensic Musicology and the Federal Rules of Evidence

The foregoing summary of basic evidence law reveals that courts should allow only portions of a typical forensic musicology opinion. As this Section will show, it is proper for musicologists to identify formal musical similarities and offer information about the musical significance and rarity of those similarities. Courts should not, however, allow experts to offer opinions about the degree of similarity shared by two works (particularly opinions about "substantial similarity") or conclusions about the existence of copying.

The potential admissibility of all evidence begins with its relevance.[253] As noted earlier, a copyright plaintiff must prove that the defendant copied from the plaintiff's work and that the defendant's copying amounted to improper appropriation.[254] Because proof of copying differs from proof of improper appropriation,[255] a musicologist could (at least in theory) provide two distinct types of relevant testimony. For copying, the musicologist would identify similarities for the jury and address the degree to which those similarities support an inference of copying. For improper appropriation, the expert would identify and evaluate aesthetic similarities by offering judgments about the degree of musical similarity created by any alleged copying. Here, the focus would be on the musical aesthetic impression made by allegedly copied features of the defendant's work.

Case law clearly states that only one of these two types of testimony is admissible. Because the test for improper appropriation depends on

---

[251] *Id.* at 154-55 (noting Carlson's "repeated reliance" on the subjective aspects of his inspection to explain his conclusion that signs of overdeflection were not substantial).

[252] *Id.* at 157.

[253] *See* FED. R. EVID. 402 (establishing that relevant evidence is admissible).

[254] *See supra* notes 42–50 and accompanying text.

[255] *See supra* notes 51–60 and accompanying text.

the perspective of a lay audience (namely an appraisal of similarity made from a lay perspective),[256] juries are well-equipped to apply the test without expert assistance. Accordingly, courts generally exclude expert testimony about improper appropriation as not helpful to the jury within the meaning of Rule 702(a).

Once again, the leading precedent is *Arnstein v. Porter*. On the issue of copying, the court wrote, "On this issue, analysis ('dissection') is relevant, and the testimony of experts may be received to aid the trier of the facts."[257] The reasoning here is simple and intuitive. Juries may not be able to analyze music well enough to identify and evaluate similarities probative of copying. If an expert can help the jury find and evaluate probative similarity, courts should allow the testimony.

By contrast, on the issue of improper appropriation, the court thought differently. The court wrote, "On that issue . . . the test is the response of the ordinary lay hearer; accordingly, on that issue, 'dissection' and expert testimony are irrelevant."[258] Once again, the reasoning is clear. If improper appropriation depends on the response of the ordinary lay hearer, then a lay jury of twelve people is well-equipped to determine the relevant response. Expert witnesses trained in music are, by definition, not laypersons. Their opinions are therefore not relevant to the issue at hand. Accordingly, expert opinions directed towards improper appropriation are not admissible because they do not help juries understand lay aesthetic responses to the two works.[259] Numerous cases follow this aspect of *Arnstein*.[260]

---

[256] *See* Arnstein v. Porter, 154 F.2d 464, 473 (2d. Cir. 1946).

[257] *Id.* at 468.

[258] *Id.*

[259] *See id.* at 473 ("The plaintiff may call witnesses whose testimony may aid the jury in reaching its conclusion as to the responses of such [lay] audiences. Expert testimony of musicians may also be received, but it will in no way be controlling on the issue of illicit copying, and should be utilized only to assist in determining the reactions of lay auditors. The impression made on the refined ears of musical experts or their views as to the musical excellence of plaintiff's or defendant's works are utterly immaterial on the issue of misappropriation; for the views of such persons are caviar to the general — and plaintiff's and defendant's compositions are not caviar.").

[260] *See* Dam Things from Denmark, a/k/a Troll Co. ApS v. Russ Berrie & Co., 290 F.3d 548, 562 (3rd Cir. 2002) (holding that "the opinions of experts may be called upon in determining whether there is sufficient similarity between the works so as to conclude

The foregoing shows that, under existing well-established doctrine, forensic musicology opinions should be used only to prove that the defendant copied from the plaintiff's work. However, it will not be enough for courts to simply insist that litigants connect forensic musicology to the issue of copying. As we will now see, questions of prejudice and reliability also affect the admissibility of forensic musicology opinions. The consequences vary with the specific content of the expert testimony in question.

1. The General Admissibility of Evidence Identifying Formal Musical Similarities

The identification of formal similarities between two works creates no meaningful evidentiary problems. Similarity clearly affects the likelihood of copying, so its basic relevance and admissibility are clear under Federal Rules of Evidence 401 and 402.[261] One might detect the possibility of prejudice or confusion because similarity also affects the existence of improper appropriation. However, this risk of prejudice is low compared to the probative value of formally identifying similarities. Identification of similarity lies at the heart of copyright infringement, and drawing attention to similarity does not require the expert to draw inferences from that similarity. Federal Rule of Evidence 403 therefore does not exclude such testimony.[262] Moreover, there is no reason to worry about exclusion based on Federal Rule of Evidence 702 because such testimony is both helpful to the jury and reliable. Musicologists certainly use terminology and notation not generally known to laypersons to identify and describe musical features. These items are highly standardized and well-understood, so there is little risk that

that the alleged infringer 'copied' the work"); Concrete Mach. Co. v. Classic Lawn Ornaments, Inc., 843 F.2d 600, 608 (1st Cir. 1988) (holding that expert testimony may aid in the assessment of "whether there are sufficient articulable similarities to justify a finding that the defendant has copied from the protected work"); Cholvin v. B. & F. Music Co., Inc., 253 F.2d 102, 103 (7th Cir. 1958) (holding that where evidence of access and similarities exist, expert testimony may be received to aid trier of facts to determine whether the similarities are sufficient to prove copying).

[261] *See* FED. R. EVID. 401–402 (defining relevant evidence as evidence tending to make more or less probable a fact of consequence to the action and establishing the admissibility of relevant evidence).

[262] *See supra* notes 177–196 (discussing effect of Rules 401 and 403).

personal fiat will overtake the provision of basic information.[263] Courts should therefore allow expert testimony identifying formal similarities between two works.

2. The More Complicated Admissibility of Expert Characterization of Identified Similarities

Concerns about reliability and prejudice become more pronounced when experts begin providing information about identified similarities to support inferences desired by the litigants. However, these concerns can be managed by insisting on proper foundations for expert testimony.

a. Expert testimony about rarity

Consider first opinions that a shared musical feature is common. Sometimes these assertions reflect common musical knowledge. For example, an expert might note that the identified similarities reflect the classical sonata form[264] or the juxtaposition of minor and major keys commonly found in blues music.[265] The expert could further contextualize such an opinion by providing additional historical or cultural background. Perhaps composers working in a certain genre or period commonly used a certain musical feature drawn from a known source. For example, classical music composers frequently use the Gregorian *Dies Irae* chant to evoke doom.[266]

Such evidence is clearly relevant because it affects the inference of copying.[267] It is helpful because a lay jury will generally not be familiar with the information. And, such testimony will be generally reliable because the expert will be able to point to scholarly publications or specific examples to establish the point. Courts should therefore

[263] *See generally* CLENDINNING & MARVIN, *supra* note 56, at 3-170 (setting forth basic principles and terminology of music theory and analysis).

[264] *Id.* at 651-60 (describing the major features of the classical sonata form).

[265] *Id.* at 708 ("Blues songs normally feature a minor-sounding vocal melody, accompanied by chords in the parallel major key.").

[266] *See infra* notes 335–336 and accompanying text.

[267] *See supra* note 158 and accompanying text.

generally admit this type of expert testimony, subject of course to cross examination.

By contrast, assertions about rarity can be evidentiarily problematic. When expert opinions about rarity do not reflect common musical knowledge, questions about reliability may emerge because the expert must rely on personal experience to make the assertion. Although Federal Rule of Evidence 702 allows experts to testify on the basis of experience,[268] the testimony must still be based on "sufficient facts or data,"[269] and there are sometimes reasons to question the sufficiency of the expert's experience. This can be seen by contrasting opinions about commonality from those about rarity.

When an expert claims that a musical feature is common, she can easily overcome concerns about reliability by offering examples of other works containing the feature in question. By contrast, when an expert asserts that a musical feature is unique or rare, he cannot provide examples that prove the claim. All he can do is state that he has studied a lot of music and has encountered this musical feature rarely, perhaps only in the plaintiff's and defendant's works.

It is highly unlikely that the expert can claim to have heard or studied every musical work. The number of extant works is far too large for anyone to have studied them all. Opinions of rarity must therefore rest on a conclusion drawn from works that the expert happens to have studied. This raises concerns about whether the works studied are sufficiently comprehensive and representative to render the expert's conclusion reliable.

Comprehensiveness matters because conclusions based on a small number of works are unlikely to be reliable. To provide an extreme example, suppose that an expert opined that the use of an organ is extremely rare in rock music on the basis of having studied five rock songs that do not feature an organ. There would be little reason to have confidence in this opinion given the huge number of rock songs that

[268] *See* FED. R. EVID. 702.

[269] *See* FED. R. EVID. 702(b).

exist, and indeed it turns out that a number of rock songs do use an organ.[270]

Representativeness also matters because different genres have different conventions. To provide another extreme example, imagine an expert who has primarily studied rock music. Imagine further that the hypothesized expert intends to testify about similarities in orchestral classical music, particularly the use of violas to carry the melody. The expert might testify that, in his experience, the use of violas to carry melodies is extremely rare. That statement might be true for rock music, but it would not be true for orchestral music.

The foregoing implies that courts should not routinely admit opinions about the rarity of musical features simply because the expert claims broad familiarity with music. Instead, the court should require the proponent to lay a foundation from which a reasonable inference of reliability can be drawn. This could mean providing examples which illustrate the expert's point, or it could mean testimony establishing that the expert has studied a reasonably comprehensive and representative set of relevant works in sufficient detail.

To see what this might mean, consider the testimony of a musicologist who testified, "I have large – you know, large files of prior art that I would investigate and compare to determine if a musical work is original in that particular element, for example."[271] The testimony implies that the expert's opinion about the prior existence of particular musical features is accurate and reliable, but the foundation is clearly inadequate. Having "large files" to investigate says nothing about the files' appropriateness, comprehensiveness, or representativeness. A court should not permit the expert to testify until the expert describes the content of the files and satisfies the court that the files offer an accurate picture of whatever "prior art" matters to the case.[272] And again, of course, such testimony would be subject to cross-examination.

---

[270] *See* Brent Chittenden, *Top Ten Organ Songs*, J. OF MUSICAL THINGS (Mar. 28, 2014), https://www.ajournalofmusicalthings.com/top-ten-organ-songs/#google_vignette [https://perma.cc/JBV6-WDSS].

[271] Testimony of Defendant's Expert Witness, Judith Finell, *supra* note 173, at 88.

[272] *See* Johannsongs-Publ'g, Ltd. v. Lovland, Nos. 20-55552, 20-55759, 2021 WL 5564626, at *1 (9th Cir. Nov. 29, 2021) (affirming trial court exclusion of expert report); Johannsongs-Publ'g, Ltd. v. Lovland, No. CV 18-10009-AB (SSx), 2020 WL 2315805, at

b. Expert opinions about aesthetic significance

A somewhat different set of concerns arise when experts testify that identified musical similarities are (or are not) aesthetically important. Here, the concerns shift to prejudice and confusion.

Consider an expert identifying formal similarities shared by two melodies and then opining that the melody of a song is aesthetically significant. Such information might be relevant to copying on the argument that a copying composer is most likely to use copied material when such material will have a meaningful effect on the quality of the composer's work. At the same time, however, opining that a similarity is aesthetically significant also addresses improper appropriation. If a similarity has a large aesthetic impact, it presumably increases the aesthetic impression of similarity, and an overall impression of aesthetic similarity forms the heart of an improper appropriation analysis.[273] This raises the possibility that a jury will consider the expert's opinion for an improper purpose.

This risk of prejudice requires courts to balance the probative value of such testimony against the risk that it will confuse and mislead the

*5 (C.D. Cal. Apr. 3, 2020) (excluding expert witness report under Rule 702 because expert did not analyze prior art before opining that certain musical features of musical works under litigation were "highly doubtful" to occur in other works); *see also* Katherine M. Leo, *Prior Art: Forensic Evidence of Musical Originality in Copyright Litigation*, 77 J. AM. MUSICOLOGICAL SOCIETY 303 (forthcoming 2024) (describing potentially arbitrary musicological choices in choosing prior art for purposes of copyright litigation).

The term "prior art" appears to be imported from patent law where prior art plays a crucial role in determining the patentability of a claimed invention. *See* 35 U.S.C. § 102(a) (conditioning patentability on "novelty, prior art"). However, the term clearly does not have the same meaning in copyright, because copyright's requirement of originality is distinct from novelty. Thus, it seems that discussions of "prior art" in copyright shed light on the possibility that a copyright plaintiff did not author portions of her own work or the likelihood that similar features arose without copying. *See* Joseph P. Fishman & Kristelia Garcia, *Authoring Prior Art*, 75 VAND. L. REV. 1159, 1160-61, 1167-70 (2022) (discussing how "prior art" is supposedly "irrelevant" in copyright law). Professors Fishman and Garcia attribute the rise of "prior art" specifically to the growing influence of expert forensic musicologists in copyright litigation. *See id.* at 1162.

[273] *See supra* notes 47–50 and accompanying text.

jury.[274] Inevitably, there will be huge variance in the probative value of these assertions and their risk of prejudice. Courts will therefore have to carefully consider the inferences to be drawn from these opinions and rule accordingly. Given that Federal Rule of Evidence 403 excludes evidence only when its probative value is substantially outweighed by the risk of prejudice,[275] it is likely that judges will find a great deal of testimony about aesthetic significance admissible. Nevertheless, it is also likely that some courts will have to exclude unusually prejudicial instances of such testimony.

3. The General Inadmissibility of Opinions About the Degree of Similarity Between Works

Things get even more challenging when experts move to opining about the degree of similarity between two works. Again, these opinions address copying and are therefore relevant. Great similarity between two works may indicate coincidence that is too large to explain except through copying. At the same time, however, these opinions clearly address improper appropriation. When a jury hears the expert's testimony that two works are very or (in some cases) substantially similar, it will understand the opinion as a conclusion that the two works are so aesthetically similar that both copying and improper appropriation have been proven.

The foregoing shows that these opinions carry a substantial risk of prejudice by improperly affecting a jury's deliberation about improper appropriation. Expert testimony about improper appropriation is clearly inadmissible under Rule 702(a).[276]Therefore, courts again must balance the probative value of such testimony against its potential for prejudice under Federal Rule of Evidence 403. This time, the balance generally tips towards exclusion.

On one hand, the probative value of testimony like this is marginal. Probative similarity and aesthetic similarity may overlap, but high degrees of aesthetic similarity do not necessarily imply probative

[274] *See* FED. R. EVID. 403 (excluding relevant evidence "if its probative value is substantially outweighed" by dangers of prejudice or confusion).

[275] *Id.*

[276] *See supra* notes 257–259 and accompanying text.

similarity because aesthetic similarities often arise for reasons other than copying.[277] Accordingly, eliciting testimony about the degree of similarity between two works means relatively little unless the expert can explain the connection between the degree of similarity and copying. Simply concluding that two works are substantially similar is not enough.

On the other hand, the prejudicial dangers of this testimony are considerable. First, as already noted, the testimony improperly invites the jury to consider expert testimony on the issue of improper appropriation.[278] Second, it confuses the jury by suggesting that aesthetic similarity alone can prove copying. If, as courts often and mistakenly state, "substantial similarity" determines both copying and improper appropriation, courts (and by extension juries) will incorrectly think that substantial aesthetic similarity can establish infringement when a determination about probative similarity must be made instead.[279] Use of terms like "very," significantly," and "substantial" draws attention to the degree of aesthetic similarity two works share, not the improbability of similarity. Jurors will then incorrectly deliberate about the existence of infringement without ever explicitly considering whether any identified similarities are sufficiently unlikely to indicate copying. Instead, they will make this decision by debating whether the two works sound too much alike (i.e., are substantially similar).

For example, consider Jury Instruction 43 from *Williams v. Gaye*:

> Extrinsic similarity is shown when two works have a similarity of ideas and expression as measured by external, objective criteria. To make this determination, you must consider the elements of each of the works and decide if they are substantially similar. This is not the same as "identical." There has been testimony and evidence presented by both sides on this issue, including by expert witnesses, as to such matters as: (a) for "Got to Give It Up" and "Blurred Lines," the so-called "Signature Phrase," hook, "Theme X," bass melodies, keyboard

[277] *See supra* notes 55–59 and accompanying text.

[278] *See supra* note 276 and accompanying text.

[279] *See supra* Part II.B.

> parts, word painting, lyrics, [and] rap v. parlando . . . . The Gaye Parties do not have to show that each of these individual elements is substantially similar, but rather that there is enough similarity between a work of the Gaye Parties and an allegedly infringing work of the Thicke Parties to comprise a substantial amount.[280]

This instruction addresses the portion of the Ninth Circuit test for infringement that comes closest to analyzing whether the defendant has copied. As noted earlier, the Ninth Circuit *Krofft* test suffers from its apparent collapse of copying and improper appropriation,[281] so instructions that embody this collapse necessarily risk confusing the jury. Here, the trial court called the jury's attention to the formal musical similarities identified and discussed by the expert witnesses who offered testimony about the degree of similarity shared by the two works. If the jury followed this instruction, it would decide the issue of copying by considering the aesthetic impression of the two works, not the likelihood of coincidence as an explanation for similarity. This represents a clear error.

When one considers the marginal value of expert testimony about the degree of similarity shared by two works against the risk that such testimony will mislead and confuse the jury, it becomes clear that any probative value is substantially outweighed by the risk of prejudice and confusion. Accordingly, under Rule 403, courts should prevent experts from testifying about the degree of similarity shared by two works.[282] Those experts should instead confine their testimony to topics like the identification of similarity, the rarity of shared features, and musical significance. Most importantly, courts should never allow experts to testify directly about whether two works are "substantially similar."

4. Opinions About the Existence of Copying and Reliability

Cumulative opinions about the existence of copying may be commonplace in copyright litigation, but they are deeply problematic.

---

[280] Williams v. Gaye, 895 F.3d 1106, 1124 (9th Cir. 2018).

[281] *See supra* Parts II.B–II.C.

[282] *See* FED. R. EVID. 403 (excluding evidence if its probative value is substantially outweighed by the risk of prejudice or confusion).

Granted, these opinions would clearly be relevant and not offered for a prejudicial purpose. However, they should be excluded for reasons of reliability under Federal Rule of Evidence 702 because the field of musicology does not give musicologists an adequate basis for rendering reliable opinions about the existence of copying. Opinions about copying therefore amount to the kind of personal say-so excluded by cases like *Kumho Tire*.

A forensic musicology opinion about copying rests on the premise that musical analysis can determine whether musical similarities result from copying or coincidence. Unfortunately, the study of music does not give musicologists this ability. It may look like musicologists can make reliable determinations analogous to those made by fingerprint, handwriting, and ballistics experts who study similarities. However, closer examination reveals that there is little reason to consider musicological conclusions about copying sufficiently reliable to satisfy Federal Rule of Evidence 702. This is true regardless of how one explains musicological expertise about copying.

a. Probabilistic inference and opinions about copying

As an initial matter, musicological opinions about the existence of copying might be reliable as intuitive but reasonably sound applications of statistical probability. It may seem correct to infer that copying has occurred because the musical features allegedly copied by the defendant are rare. However, as Professors Buccafusco and Tushnet explain, such reasoning violates basic principles of probabilistic reasoning.[283] Error arises because opinions based solely on the rarity of particular musical similarities improperly disregard base rate probabilities associated with potential causes for unlikely sets of musical similarities. This makes it impossible to reliably determine whether copying or coincidence caused a given instance of potential copying.[284]

Suspicious similarities (i.e., similarities that reflect potential copying) have two possible causes. Either the similarities arose because the defendant copied from the plaintiff work, or the defendant did not copy

[283] *See* Buccafusco & Tushnet, *supra* note 25, at 236-37.

[284] *See infra* notes 286–291 and accompanying text.

from the plaintiff's work and the similarities arose from coincidence.[285] If one composer is copying from another, suspicious similarities would likely occur. If one composer is not copying from another, suspicious similarities might be unlikely to follow, but the chance is not zero. Accordingly, one cannot determine whether copying or coincidence is a better explanation for a given instance of suspicious similarity simply by establishing that a single act of composition would be highly unlikely to create the similarities in question.

To see this, one must realize that a probabilistic inference of copying depends on predictions about a set of musical similarities whose origins differ. Over time, an enormous number of composers will create musical works. The overwhelming majority of them will not copy from the plaintiff's work. One may hypothesize that a small minority of them will. Some very small percentage of the composers working independently will coincidentally author works bearing suspicious similarity to the plaintiff's work. A much larger percentage of the composers who copy will also create works with suspicious similarity. When a musicologist encounters a work bearing suspicious similarity to the plaintiff's work, she must decide which type of composer created the similarities. She cannot do that without knowing the number of independently working and copying composers and how easily suspicious similarity arises.[286]

To build upon a hypothetical given by Buccafusco and Tushnet,[287] let us hypothesize that two works share a suspiciously unlikely set of musical similarities. Let us assume further that there are 1,000,000 compositions created by composers who do not copy from the plaintiff and that the likelihood of coincidental suspicious similarity is 1/1,000 (.001). At the same time, let us imagine that there are 100 compositions created by composers who copy from that source, and that the likelihood of suspicious similarity is 9/10 (0.9 — not 100% to account for the possibility that a copier will successfully hide her copying).

[285] For purposes of analytical simplicity, I will include within coincidence the possibility that the defendant drew from a source other than the plaintiff's work.

[286] *See* Buccafusco & Tushnet, *supra* note 25, at 253.

[287] *Id.* at 247-48 (illustrating how base rates greatly affect the statistical likelihood of negligence).

Under these facts, we would expect one thousand instances of coincidental suspicious similarity[288] and ninety instances of suspicious similarity created by copying.[289] Because there are more instances of suspicious similarity created coincidentally than those created by copying, any given occurrence of suspicious similarity would be more likely to have arisen from coincidence than copying.[290] This example

[288] The calculation would be: (1,000,000 compositions) x (.001 likelihood of suspicious similarity) = 1,000 expected instances of suspicious similarity. *See id.* at 248 (performing similar calculation in the context of a tort hypothetical).

[289] The calculation would be: (100 compositions) x (0.9 likelihood of suspicious similarity) = 90 expected instances of suspicious similarity.

[290] The probability of copying given the existence of suspicious similarity is 90/1,090. Intuitively, the hypothetical creates 1,000 expected instances of coincidental suspicious similarity and 90 instances of suspicious similarity arising from copying. Thus, the probability of suspicious similarity from copying is 90/1,090.

More formally, the use of Bayes' Theorem reaches the same result as follows:

Let:

A = composition involves copying

B = suspicious similarity

We may compute Probability of A, Probability of B, and Probability of A and B as follows:

P(A) = 100/1,000,100 (100 instances of copying out of 1,000,100 total compositions)

P(B) = 1,090/1,000,100 (1,090 instances of suspicious similarity out of 1,000,100 compositions)

P(B|A) = 9/10 (given by the hypothetical)

Bayes' Theorem:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Or:

Probability of A given B = (Probability of B given A) * (Probability of A)/ (Probability of B)

Substituting in values computed above:

Probability of A given B = (9/10)*(100/1,000,100)/(1,090/1,000,100)

After reducing the right side of the equation, the solution appears:

Probability of A given B = 90/1,090

demonstrates that a low likelihood of coincidental similarity alone cannot support an inference of copying as a matter of probability. Appropriate inferences about copying require information that musicologists do not have. We do not know the relative prevalence of certain compositional practices (copying versus not copying from the plaintiff's work), nor do we know the likelihood that those practices will lead to similarities suspicious of copying. Accordingly, under the current state of available information, one cannot use probability to support the accuracy or reliability of expert opinions about the existence of copying.[291]

b. Experience and opinions about copying

Alternatively, one might argue that musicological opinions about copying gain reliability as the product of experience acquired through the study of music. A musicologist could claim that she has learned to recognize the difference between suspicious coincidental similarity and copying through repeated exposure to appropriate exemplars. A musicologist can therefore tell whether suspicious similarities reflect coincidence or copying, just as a sommelier can, at least occasionally, recognize by blind tasting the varieties of grape or country of origin associated with various wines.[292] However, although superficially plausible, such a claim falls apart because the appropriate experiential background does not exist.

A sommelier's ability to identify and distinguish wines arises from a very particular experience that requires exemplars of established authenticity.[293] Over time, the sommelier encounters and studies these

See QUANTUM SCIENTIFIC PUBLISHING, PROBABILITY 45-47 (2017) (explaining conditional probability and Bayes' Theorem); Chris F. Westbury, *Bayes' Rule for Clinicians: An Introduction*, 1 FRONTIERS IN PSYCH. 1, 2-3 (2010), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3153801/ [https://perma.cc/A9JP-479P] (explaining Bayes' Theorem).

[291] *See* Buccafusco & Tushnet, *supra* note 25, at 246-53 (explaining why information about base rates is necessary for sound probabilistic calculation).

[292] *See* THE OXFORD COMPANION TO WINE 83 (Jancis Robinson ed., 2d ed. 1999) (describing blind tasting as "a particularly masochistic but potentially rewarding exercise" and noting difficulties encountered by "trained tasters" in correctly identifying variety of grape and vintage).

[293] *See* Kendeigh, *Build a Blind Tasting Collection for Wine Study!*, THE GRAPE GRIND, https://thegrapegrind.com/blindtastingsystem/ (last visited July 24, 2024)

exemplars, eventually learning the characteristics of various grapes and countries of origin.[294]

Musicological study does not provide this kind of experience because the necessary exemplars for comparison do not exist. To be sure, known instances of copying exist because composers sometimes explain what they have done.[295] However, in order for our hypothetical musicologist

[https://perma.cc/H28E-D52P] (offering a system for assembling a large collection of wine to improve blind tasting skills).

[294] Certification tests for sommeliers include blind tastings that ask the exam taker to identify the variety of grape and country of origin associated with particular wines. *See Certified Sommelier Examination*, CT. OF MASTER SOMMELIERS AMS., https://www.mastersommeliers.org/wp-content/uploads/2023/12/CMS-A-Certified-Sommelier-Deductive-Tasting-Grid.pdf (last visited July 24, 2024) [https://perma.cc/6AKU-SC9J]. Gaining the ability to successfully take this sort of exam requires, among other things, time spent tasting and studying wine. The Court of Master Sommeliers offers a workshop that offers participants the chance to taste "24 examples of classic wines from around the world." *See Deductive Tasting Method Workshop*, CT. OF MASTER SOMMELIERS AMS., https://www.mastersommeliers.org/certification/courses/the-deductive-tasting-method-workshop/ (last visited July 24, 2024) [https://perma.cc/6CYV-THAQ]. Beyond this, the aspiring sommelier must spend time tasting a great deal of wine. *See* Vicki Denig, *15 Sommelier-Level Moves for Learning About Wine*, FOOD & WINE, https://www.foodandwine.com/wine/learn-about-wine (last updated Oct. 28, 2018) [https://perma.cc/8UUP-NNT8] ("Most sommeliers agree that tasting as much as possible is the best way to learn about wine."); *Journey to Certified Sommelier*, THE GRAPE GRIND, https://thegrapegrind.com/my-journey-to-certified-sommelier/ (last visited July 24, 2024) [https://perma.cc/PA8J-4H9A] (describing how one certified sommelier studied wines for 2-3 hours per day).

[295] For example, a composer writing a set of variations based on a theme often identifies the borrowed theme. *See* JOHANNES BRAHMS, VARIATIONS ON A THEME BY HAYDN, OP. 56 (1873); SERGEI RACHMANINOFF, RHAPSODY ON A THEME OF PAGANINI, OP. 43 (1934). In other cases, a composer may leave behind writings that explain what has happened. *See also* SERGEI BERTENSSON & JAY LEYDA, SERGEI RACHMANINOFF: A LIFETIME IN MUSIC 33 (1956) (quoting SERGEI RACHMANINOFF, RHAPSODY ON A THEME OF PAGANINI, OP. 43 (1934) ("All variations of the Dies Irae would be for the evil spirit.")).

Although instances of admitted copying provide exemplars for the experience claimed by our hypothetical forensic musicologist, it is worth noting that their value is marginal. A composer who admits copying is not trying to pass someone else's work off as his own. He wants the audience to know what he has done. By contrast, the typical music copyright defendant stands accused of copying from the plaintiff while "covering his tracks" in hopes that no one will notice. It is exceedingly difficult, if not impossible, to collect verified exemplars of this sort of copying because no means of verification exists. By definition, the composer will not admit to his behavior. In theory, a composer covering his tracks might secretly admit to his behavior, or perhaps a witness exists.

to develop the necessary experience, she must compare known instances of copying to known instances of suspicious similarities that verifiably arose from coincidence. Unfortunately, it is effectively impossible to identify a sufficient number of works with coincidental suspicious similarities to provide the necessary experience.

A suspicious similarity is one that might or might not have arisen from copying. If such similarity resulted from coincidence (as indeed it must for the discussion at hand), the composer who wrote the work would have no awareness of copying and would see no need to declare the work original. At a later date, others might notice the suspicious similarity and raise the possibility of copying. The composer could respond with denials of copying, but her credibility would be shaky because the denials, if given, would be self-serving assertions advanced to protect the composer's reputation. Only in the rarest of circumstances could a composer have incontrovertible proof against copying. For example, a composer might try to disprove copying by displaying drafts of her work. Such evidence would constitute plausible but not incontrovertible proof against copying because drafts cannot prove where a suspicious similarity comes from. Accordingly, it is hard to imagine that a musicologist could ever credibly claim to have studied enough verified exemplars to develop the ability to distinguish copying from coincidentally created suspicious similarity.

c. Musicological methods and opinions about copying

Finally, a musicologist could claim that her ability to distinguish copying from coincidence comes not from statistics or the study of exemplars but from the skilled application of musicological methods that reliably identify copying. This claim is also superficially plausible but questionable. To be sure, musicologists often use musical analysis to assert that one composer has copied from another.[296] However, musicologists do not support these claims with a method that reliably identifies copying. Instead, musicologists use whatever arguments are

Even so, the number of such instances is probably insufficiently large to provide the kind of experience at issue here.

[296] *See infra* notes 299, 362 and accompanying text (discussing various instances of musicologists analyzing whether one composer has copied from another).

intuitively handy, frequently offering nothing more than the identification of similarity and an assertion that copying has taken place.[297] These musicological opinions about copying therefore amount to educated personal hunches. Nothing tells a musicologist which similarities matter or the degree and number of similarities necessary to establish copying.[298]

To see this, let us examine claims about copying made by academic musicologists, a group whose members educate and sometimes serve as forensic musicologists. If there is a way to reliably identify copying, academic musicologists would presumably use it. However, an examination of scholarship about copying reveals no such thing. Musical analysis can identify clues that suggest copying, but it cannot establish when clues amount to proof. Conclusions about copying therefore depend heavily on personal fiat.

For example, consider Jeremy Orosz's interesting and informative article *"Can't Touch Me": Television Cartoons and the Paraphrase of Popular Music*.[299] In it, he makes the claim that composers writing music for television cartoon shows like *The Simpsons* and *Family Guy* "craft copyright-dodging sound-alike versions of songs that are canny enough to pass as the tunes upon which they are based."[300] Doing this allows the

[297] *See id.*

[298] I should make clear that I am not criticizing musicologists for their practices. Nor am I suggesting that the field of musicology lacks value. Musicologists make assertions about copying to explore possible influences in a composer's work and to understand how one composer's work fits with others. *See* Burkholder, *Musical Borrowing*, *supra* note 7, at 223-26 (identifying the study of musical borrowing as a field and describing why musicologists would study it); Burkholder, *Uses of Existing Music*, *supra* note 24, at 851-52 (same); *see also* JUSTIN A. WILLIAMS, RHYMIN' AND STEALIN': MUSICAL BORROWING IN HIP HOP 1 (2013) (basing the author's explanation of hip-hop on "a crucial premise: the fundamental element of hip-hop culture and aesthetics is the overt use of preexisting material to new ends"). It may be entirely appropriate to consider assertions of borrowing without reliable methods of proof when the objective is musical criticism. Even hypothesizing that one composer has borrowed from another opens the door to valuable study of how that composer's work relates to other works. This does not mean, however, that musicological proof, whatever it is, satisfies the evidentiary standards of the legal system.

[299] Jeremy W. Orosz, *"Can't Touch Me": Television Cartoons and the Paraphrase of Popular Music*, 33 CONTEMP. MUSIC REV. 223 (2014).

[300] *Id.* at 224.

show to evoke, perhaps for humorous purposes, associations with the earlier musical work.[301]

Orosz provides his reader with many examples of supposed copying, including one in which he claims that *The Simpsons*' composer copies from *The Little Mermaid* song "Under the Sea."[302] As proof, Orosz offers a side-by-side transcription of the melodies from the two songs, along with a brief analysis noting similarities in rhythm, instrumentation, and timbre.[303] Orosz repeats this method of proof numerous times for other examples that he gives.[304]

Orosz's approach illustrates the fundamental problem identified here. To the extent that reliability stands on methods consistently applied by educated and trained musicologists, no foundation exists. Orosz identifies similarities, posits a motive for copying, and concludes that copying exists.[305] Orosz may be right.[306] However, if his reader disagrees, Orosz invokes no standard principles, analyses, or methods that would settle the disagreement or even add additional persuasion. If there are ways to use musical analysis to prove copying, Orosz does not invoke them. A skeptic would recognize that a reasonable suspicion of copying exists, but nothing prevents her from chalking up the similarities to coincidence and the use of conventions associated with calypso music.[307] Accordingly, sensibilities peculiar to the analyst control the outcome of the inquiry.

---

[301] *Id.* at 223-24.

[302] *Id.* at 224.

[303] *Id.* at 224-25.

[304] *See id. passim*.

[305] *See id.*at 224 (stating that composers "craft copyright-dodging sound-alike versions of songs that are canny enough to pass as the tunes upon which they are based").

[306] One can hear the songs in question on YouTube at: DisneyMusicVEVO, *The Little Mermaid — Under the Sea (from The Little Mermaid) (Official Video)*, YOUTUBE (Nov. 22, 2013), https://www.youtube.com/watch?v=GC_mV1IpjWA [https://perma.cc/R6LP-8XXC] ("Under the Sea" from *The Little Mermaid*); Omar PhD, *Homer Under the Sea*, YOUTUBE (Nov. 20, 2010) https://www.youtube.com/watch?v=twVvnr7bOw4 [https://perma.cc/DW6Z-PUC7] (from *The Simpsons: Homer Badman* (20th Century Fox television broadcast Nov. 27, 1994)).

[307] *See infra* notes 347, 362 and accompanying text (describing how disagreements between academic musicologists turn largely on personal opinion not well-connected to any methodology).

Similar problems exist in writings about the extent to which the famous composer Gustav Mahler based his work on that of his little-known contemporary Hans Rott. Debate began after Paul Banks wrote about Rott as a composer who had survived "in the penumbra of music history" largely because Mahler and Anton Bruckner (a famous composer and their mutual teacher) admired him.[308] Banks also wrote that Mahler greatly admired Rott's "Symphony in E Major," "incorporating into [Mahler's] Second, Third, Fifth and Seventh Symphonies explicit references to material from his friend's magnum opus."[309]

Banks did not allege that Mahler had copied to accuse Mahler of plagiarism. Instead, he argued that Mahler's apparent use of material from Rott was "an expression of grief" about Rott's early death[310] and that Rott was a composer of greatly underappreciated talent.[311] However, others raised the stakes, arguing that the alleged copying diminished Mahler's artistic reputation.[312] Noted conductor Paavo Jarvi stated that Mahler's use of Rott amounted to "a historic crime . . . a historic case of plagiarism, today Mahler would be sued for

[308] Paul Banks, *Hans Rott, 1858–1884*, 125 MUSICAL TIMES 493, 493 (1984).

[309] *Id.* at 494. Banks wrote again about "what appear to be explicit borrowings of material and ideas" from Rott by Mahler in 1989. *See* Paul Banks, *Hans Rott and the New Symphony*, 130 MUSICAL TIMES 142, 142 (1989).

[310] Banks, *supra* note 308, at 494.

[311] *Id.* at 495.

[312] *See* Helmuth Kreysing & Frank Litterscheid, *Mehr als Mahlers Nullte!*, *in* MUSIK-KONZETPE 91 MAHLER DER UNBEKANNTE BEKANNTE 46, 46 (Heinz-Klaus Metzger & Rainer Riehn eds., 1996) (similarity between Mahler's works and those of Rott "questions the scope of originality of Mahlerian symphonies"). Translation of the original German text is from a PhD thesis. ROGIER VAN GULICK, GUSTAV MAHLER AND HANS ROTT: THE MUSICOLOGICAL ASSESSMENT OF A RELATIONSHIP, UNIVERSITEIT UTRECHT (2017), https://studenttheses.uu.nl/bitstream/handle/20.500.12932/28926/Thesis%20Rogier%20van%20Gulick.pdf?sequence=1&isAllowed=y [https://perma.cc/8ZCR-R6P2].

plagiarism."[313] Others have disagreed, finding coincidence a plausible explanation for many identified similarities.[314]

Conclusions about the size of Mahler's debt to Rott (if any) depend on how much Mahler copied from Rott. If musicology offered a methodology that could reliably identify or debunk instances of copying, one would expect those debating Mahler's debt to invoke that methodology. Yet the relevant authors do no such thing. It is apparently enough to identify similarities and assert the existence or non-existence of copying without reference to accepted methods of proof.[315]

J. Peter Burkholder recognized this problem in his article *Musical Borrowing or Curious Coincidence?: Testing the Evidence*.[316] Like all musicologists, he takes it as an article of faith that copying occurs throughout composition.[317] Yet, although numerous scholarly writings address the issue of copying,[318] he asks "[h]ow can we be sure that a similarity between two pieces results from borrowing and is not a coincidence, or the result of drawing on a shared fund of musical ideas?"[319]

---

[313] Kenneth Woods, *A flawed perspective? Mahler and Rott*, KENNETH WOODS: BLOG (Apr. 27, 2010), https://kennethwoods.net/blog1/2010/04/27/a-flawed-perspective-mahler-and-rott/ [https://perma.cc/84C6-6QC5]. An audio recording of Jarvi making these statements is available at https://ionarts.blogspot.com/2010/04/paavo-jarvi-on-hans-rott.html [https://perma.cc/UVQ6-HCNY].

[314] *See* Stephen McClatchie, *Hans Rott, Gustav Mahler and the "New Symphony": New Evidence for a Pressing Question*, 81 MUSIC & LETTERS 392, 401 (2000) (arguing that "it is just as plausible that [identified similarities] originated independently from 'the same soil'. . . as Mahler put it").

[315] *See* Banks, *supra* note 308, at 494 (identifying apparent copying of Rott by Mahler); *see also supra* note 54 (listing articles containing assertions of copying by Debussy, all without reference to an accepted method of proof); *infra* notes 338-339 (discussing assertions about musical borrowing by Brahms, all without reference to an accepted method of proof).

[316] Burkholder, *Musical Borrowing*, *supra* note 7, at 223.

[317] *Id.* at 223 ("Many pieces of music borrow from other pieces in some way . . . .").

[318] He cites to a database of scholarship about borrowing. *Musical Borrowing & Reworking: An Annotated Bibliography*, CTR. FOR THE HIST. OF MUSIC THEORY AND LITERATURE https://chmtl.indiana.edu/borrowing/ (last visited Sept. 20, 2024) [https://perma.cc/9LE2-63KN].

[319] Burkholder, *Musical Borrowing*, *supra* note 7, at 223.

Burkholder handles this problem by organizing "the types of evidence scholars have used to advance or refute claims of borrowing" into three categories: analytical evidence, biographical and historical evidence, and the musical purpose of any alleged borrowing.[320] He discusses all three to argue that conclusions about potential copying improve when analysts address all three types of evidence.[321]

Burkholder's treatment of musical analysis parallels what expert witnesses attempt to do. He writes, "[S]cholars try to show a similarity between two pieces that is so strong that borrowing is the best explanation — or, conversely, that the similarity is not strong enough to justify such a claim."[322] He then explains that as the extent of similarity between two works increases, the case for copying strengthens.[323] In so doing, he criticizes some musicologists for concluding that copying exists on the basis of too few similarities.[324]Burkholder goes on to describe how the exactness of a match between similarities, the number of shared elements, and the rarity of shared elements affect the likelihood of copying. For all of these things, he makes the commonsense claim that as these things increase, so does the strength of any case for copying.[325]

Burkholder makes similar claims about evidence derived from sources other than musical analysis. Obviously, a composer cannot copy without being familiar with the alleged source material. Burkholder cautions us that relying on musical analysis can "lead us astray."[326] For example, he describes how similarities between Claude Debussy's "Pour la danseuse aux crotales" and Alban Berg's "Warm die Lufte" are, in his opinion, "so striking that it is hard to imagine two composers arriving at the same idea independently."[327] Nevertheless, he believes that such a conclusion would be wrong. Berg was an admirer of Debussy, but his work predated Debussy's, making it impossible for him to be a copier. While perhaps

[320] *Id.* at 224.

[321] *Id.* at 226-27.

[322] *Id.* at 227.

[323] *Id.* at 227-28.

[324] *Id.* at 228.

[325] *Id.* at 228-36.

[326] *Id.* at 237.

[327] *Id.* at 237.

Debussy copied from Berg, Berg was not well known in France at the time, and there is no evidence that Debussy had seen or heard Berg's work. Burkholder then states his belief that both composers probably borrowed from Ravel's *Gaspard de la nuit*. He describes similarities that both works share with *Gaspard*, including similarities that Debussy and Berg's works do not share with one another.[328]

Burkholder finishes his essay by noting the importance of motive, writing that "proof of borrowing is incomplete until its *purpose* is clear."[329] This again follows common sense. It is easier to believe that a composer borrowed if there was a reason for her to have done so. Otherwise, if no explanation for copying exists, the possibility of coincidence seems more plausible.[330] For example, composers sometimes borrow for structural or thematic reasons, to emulate or evoke the work of earlier well-known composers, to make extra-musical associations, or for humor.[331] Interestingly, Burkholder claims that questions of purpose rarely affect legal cases beyond the fair use analysis of parody because the legal system generally assumes that profit motives explain why a defendant would want to copy.[332]

Burkholder's effort is about as close as one will come to a musicological method for proving copying. Burkholder demonstrates that careful attention to multiple angles improves opinions about copying. Nevertheless, Burkholder still fails to bring reliability to the enterprise. He demonstrates the characteristics of stronger and weaker cases for copying, but he never indicates how to know when the case has been made, other than to label some cases beyond doubt.[333] Someone applying Burkholder's methods may become more or less suspicious of copying, but the ultimate conclusion rests on the individual's personal predisposition to believe that copying has taken place. Indeed, some of Burkholder's own examples prove this very point.

---

[328] *Id.* at 237-41.

[329] *Id.* at 247.

[330] *Id.*

[331] *Id.* at 247-50.

[332] *Id.* at 226.

[333] *Id.* at 227 (stating "there can be no doubt" about two instances of claimed borrowing).

Consider again Burkholder's writing about the musical analysis of similarity, the backbone of any forensic musicologist's opinion about copying. To illustrate how a musicologist should handle musical similarity, Burkholder discusses various instances in which composers may have copied the Gregorian *Dies Irae* chant.[334] For those unfamiliar with it, the opening notes of *Dies Irae* (Latin for "day of wrath") have become a musical standard for evoking doom.[335] Many composers are thought to have deliberately copied ("quoted") *Dies Irae*, especially its first four notes, for this purpose.[336]

Burkholder discusses three instances of possible *Dies Irae* copying advanced by musicologists: Berlioz's *Symphonie Fantastique*, Liszt's *Totentanz*, and Brahms' *Intermezzo in E-flat Minor, op. 118/6*. He considers the first two obvious instances of borrowing because the similarities of melody are exact and extend for nineteen notes.[337] By contrast, he considers the Brahms case doubtful because the identical sequence spans only four notes[338] — even though composers often appear to quote only those four notes.[339]

---

[334] *Id.*

[335] The chant may be heard on YouTube. Tu Es Petrus, *Dies Irae — Gregorian Chant (with lyrics and translation)*, YOUTUBE (Sept. 4, 2021), https://www.youtube.com/watch?v=2OBB5-bP6qs [https://perma.cc/PC4S-KVXH].

[336] *See* Martin Hoondert, *Dies Irae: Origin and Appropriations*, *in* MUSIC IN THE APOCALYPTIC MODE 44-64 (Lorenzo DiTommaso & Colin McAllister eds., 2023). A particularly entertaining compilation of supposed invocations is the apparent use of *Dies Irae* in movie music. Brian LaGuardia, *Dies Irae in the Movies*, YOUTUBE (Apr. 5, 2013) https://www.youtube.com/watch?v=0hL1m4hGBVY [https://perma.cc/Z82V-NTVR].

[337] Burkholder, *Musical Borrowing*, *supra* note 7, at 227-28.

[338] *Id.* at 228-29.

[339] For assertions that Brahms copied the *Dies Irae* theme in his *Intermezzo in E-flat Minor*, see Nicole Grimes, *Brahms as a Vanishing Point in the Music of Wolfgang Rihm: Reflections of Klavierstuck Nr. 6*, *in* MUSIC PREFERRED: ESSAYS IN MUSICOLOGY, CULTURAL HISTORY, AND ANALYSIS IN HONOUR OF HARRY WHITE 523, 541 (Lorraine Byrne Bodly ed., 2018) (asserting that "[t]he entire piece is built upon an obsession with a single motive: the archaic *Dies Irae* melody"); David Lidov, *Reflections on Musical Topics and Musical Character in Performance*, *in* MUSIC SEMIOTICS: A NETWORK OF SIGNIFICATIONS, 163, 168 (Esti Sheinberg ed., 2012) (stating that the *Intermezzo* "begins with an invocation of *Dies Irae* . . . which wants only the interpolation of one D♭"); JOHN BELL YOUNG, BRAHMS: A LISTENER'S GUIDE 80 (2008) (referring to the "wholesale appropriation of Dies Irae").

As a matter of common sense, Burkholder may be right. Longer sequences of similarity generally raise stronger inferences of copying. But when is a sequence long enough to prove copying? Burkholder does not say, and indeed he cannot. Granted, Burkholder suggests that an identical nineteen note melodic sequence is clearly copying because "[t]he likelihood of generating two identical nineteen-note diatonic melodies by random processes is vanishingly small, less than 1 in 10 quadrillion."[340] However, it is incorrect to understand Burkholder's point as a blunt assertion that this sort of calculation can settle the question of copying.

As a matter of statistics, all of the objections raised by Professors Buccafusco and Tushnet still apply. Moreover, Burkholder's calculation represents the number of possible nineteen-note sequences that can be randomly generated within a seven-note diatonic set.[341] While the number of possible sequences is undoubtedly extremely large, one must be careful about taking Burkholder's calculation and assertion literally because not all sequences are equally likely to be chosen. Playing random notes on a piano usually does not create a comprehensible (let alone pleasing) tune because relatively few sequences of notes make musical sense. Composers therefore choose notes that fit the principles behind this sense, making similarity considerably more likely than one might otherwise expect.[342]

Furthermore, one should appreciate that, under the law of large numbers, very unlikely events become likely to occur when enough trials take place. To see this, consider an outcome (perhaps the random selection of a certain sequence of notes) that has a .00001% (1 in 100,000) probability of occurrence. If only one composition is written (i.e., only one trial takes place), the likelihood of the sequence occurring is 1 in 100,000. But as more and more compositions get written, the chance of the sequence being chosen at least once begins to rise to the point that one would expect the sequence to occur. To illustrate how

[340] Burkholder, *Musical Borrowing*, *supra* note 7, at 227.

[341] *Id.* at 227 n.7. A seven-note diatonic set represents the notes on a scale. *See* CLENDINNING & MARVIN, *supra* note 56, at 50 (explaining chromatic and diatonic collections of notes).

[342] *See* CLENDINNING & MARVIN, *supra* note 56, at 177 (setting forth guidelines for writing a good melodic line).

quickly this expectation arises, one can calculate how many trials are necessary before the probability of an unlikely event occurring reaches 50% when each trial is independent of the others. In the case of a .00001 (1 in 100,000) likelihood event, the number of trials needed is only 69,266.[343] Thus, even though the odds of random duplication may be quite long, the enormous amount of music composed over years (even centuries)[344] makes it rather likely that many seemingly unique musical features will be independently replicated, and probably multiple times. An analyst considering an instance of possible copying must therefore decide whether he is looking at an unlikely instance of actual copying or a coincidence that is very likely to have occurred.

This point is implicit in Burkholder's analysis. In his discussion of non-musical evidence, Burkholder states that the similarities between Berg and Debussy's works are "so striking that it is hard to imagine two composers arriving at the same idea independently," but he ultimately concludes that neither copied from the other.[345] He understands that musical analysis alone provides important but unreliable clues about the existence of copying. Musical analysis contributes evidence that makes a claim of copying stronger or weaker, but it alone cannot tell us when the case for copying has been made. Other information has a vital role to play, and even then, the ultimate conclusion of copying requires a personal judgment that heavily depends on idiosyncratic choices about what evidence to consider and predispositions to believe that copying explains similarity. Those who disagree with Burkholder about Brahms'

[343] *See* H.B. Berman, *Binomial Probability Calculator*, STAT TREK (2024), https://stattrek.com/online-calculator/binomial [https://perma.cc/8XDD-HJUM].

[344] Industry publications estimate that 100,000 tracks get added to streaming services every day. Even if some of those tracks represent recordings of the same composition, over years, the number of new compositions will be enormous. *See* Tim Ingham, *It's Happened: 100,000 Tracks Are Now Being Uploaded to Streaming Services Like Spotify Each Day*, MUSIC BUS. WORLDWIDE (Oct. 6, 2022), https://www.musicbusinessworldwide.com/its-happened-100000-tracks-are-now-being-uploaded/ [https://perma.cc/WV8P-TD53]; Glenn Peoples, *The Ledger: Are There Really 100,000 New Songs Uploaded a Day? Maybe More.*, BILLBOARD (Feb. 2, 2023), https://www.billboard.com/pro/how-much-music-added-spotify-streaming-services-daily/ [https://perma.cc/7XS2-ZDD5].

[345] Burkholder, *Musical Borrowing*, *supra* note 7, at 237; *see id.* at 237-41.

potential copying of *Dies Irae* are not demonstrably wrong. They are simply more easily convinced of copying than Burkholder.[346]

For a concrete example of this, consider Carol Baron's review[347] of Burkholder's book *All Made of Tunes: Charles Ives and the Uses of Musical Borrowing*.[348] Burkholder uses this book to catalog and discuss exhaustively the many apparent borrowings found in Ives' work, separating them into fourteen different categories.[349] Banon compliments Burkholder,[350] but this does not mean that she accepts all of Burkholder's assertions about copying. In particular, she criticizes Burkholder for basing his analysis too heavily upon the melodic aspect of Ives's music and "identified" fragments (i.e., portions of melodies that Burkholder believes were copied from other works).[351] According to Baron, "when analyses of other parameters are introduced, the conclusions reached may be contraindicative . . . ."[352] Later, she writes, "Another problem arising from Burkholder's overemphasis on the process of borrowing is that of ambiguous or even incorrect identification."[353] She then goes on to argue that, for example, Burkholder is wrong to claim that Ives based the slow movement theme

[346] Interestingly and tellingly, none of those who apparently believe that Brahms copied the *Dies Irae* offers any proof that he or she is correct. *See supra* note 339. Again, musicologists appear willing to identify similarity and proceed on the belief that the copying is self-evident. Obviously, that must not genuinely be the case given Burkholder's disagreement. To drive the point home one more time, it is instructive to discover that not every musicologist considers even Debussy's widely mentioned borrowing from Wagner in "Golliwog's Cakewalk" to have been proven. *See* DeVoto, *supra* note 54, at 146 (stating that widely accepted borrowing from Wagner by Debussy in "Golliwog's Cakewalk" "may well be merely accidental").

[347] Carol K. Baron, *Review*, 53 J. AM. MUSICOLOGICAL SOC'Y 437 (2000).

[348] J. PETER BURKHOLDER, ALL MADE OF TUNES: CHARLES IVES AND THE USES OF MUSICAL BORROWING (1995).

[349] *Id.* at 3-4.

[350] Baron, *supra* note 347, at 437 (congratulating Burkholder for "the first comprehensive study of Ives's music").

[351] *Id.* at 437-38.

[352] *Id.* at 438.

[353] *Id.* at 439.

from his First Symphony on the slow movement from Dvorak's Ninth Symphony (commonly known as the "New World Symphony").[354]

Baron's reaction to Burkholder again demonstrates that musicologists do not have a method to determine reliably when copying has occurred. They may analyze the same works, but their methods of analysis differ for personally subjective reasons. Burkholder gives considerable weight to relatively isolated and brief instances of formal musical similarity, leading him to conclude that Ives copied from Dvorak. He spends over four pages exhaustively analyzing similarities and differences between the two works, ultimately characterizing Ives' work as an "elegant condensation" of Dvorak's theme that simultaneously pays homage to and challenges Dvorak.[355] Baron disagrees, stating that Burkholder attaches too much significance to minor similarities while criticizing him for ignoring or overlooking other aspects of the works in question. In her opinion, the fragments identified by Burkholder are too brief to be convincing evidence of copying, making coincidence the likely explanation.[356] Tellingly, neither musicologist argues that musicological principles establish the superiority of his or her analysis. Instead, each simply emphasizes musical features that support his or her personal intuition about the existence of copying.

Observations like this lead some to argue that musicological "proof" of copying amounts to nothing more than rank speculation. For example, in his article *On Influence and Borrowing*,[357] W. McNaught criticizes the musicological attribution of similarities to copying or

[354] *Id.* at 439-40. Baron is not the only reviewer who is not entirely convinced of Burkholder's accuracy. *See e.g.*, Burton W. Peretti, *All Made of Tunes: Composers, Music and American Culture*, 38 AM. STUD. 139, 142 (1997) (stating that Burkholder's analysis "will not settle all arguments" about Ives's alleged borrowing and that the author is "certain" that Burkholder misidentified a claimed borrowing).

[355] BURKHOLDER, *supra* note 348, at 89-92 (presenting claim that Ives copied from Dvorak).

[356] Baron, *supra* note 347, at 439-40 (criticizing Burkholder's inattention to "scalar placement," "harmonic congruence," "melodic contractions and expansions," and "displaced rhythms and long stretches of unrelated filler" while noting that fragments of the sort identified by Burkholder "could be found in any number of diatonic pieces").

[357] W. McNaught, *On Influence and Borrowing*, 90 MUSICAL TIMES 41 (1949).

coincidence as "arbitrary," "unscientific," and "unstable."[358] He points out that composers necessarily create within musical traditions and conventions.[359] Composers therefore inevitably select musical phrases that others have chosen, and musicologists are often too eager to overlook this likelihood.[360] McNaught concedes (and indeed believes) that composers copy from others often, so he is not unhappy with musicologists considering the possibility of copying and its implications.[361] He is, however, making the point that musical analysis is not capable of reliably separating coincidental similarity from copying.[362]

d. Opinions about copying and the Federal Rules of Evidence

The foregoing shows that forensic musicology opinions about the existence of copying do not meet the requirement of reliability embodied in Federal Rule of Evidence 702. To see this, consider how the testimony of forensic musicologists suffers from the same shortcomings as the expert's testimony in *Kumho Tire*. In *Kumho Tire*, the expert Carlson examined physical evidence of potential tire wear and opined

[358] *Id.* at 45; *see also* Durand R. Begault, Heather D. Heise & Christopher A. Peltier, Forensic Musicology — An Overview 1-2 (unpublished manuscript), https://www.jurispro.com/files/articles/forensicmusicology_6302.pdf [https://perma.cc/DRP8-4T72] (identifying "disparaged" analytical methods employed by forensic musicologists as unacceptably subjective).

[359] McNaught, *supra* note 357, at 44 ("Throughout the classical-romantic period the ordinary vocabulary of music had in it a thousand and one short phrases and shapes that amounted to little in themselves but could act as units in a fabric of significant melody. All day long composers were borrowing these bits and pieces, not from each other, but from the daily language of their art. This is of course well understood by anybody who takes an interest in the fabric of music.").

[360] *Id.* at 41 (taking the position that coincidental similarity is far more common than musicologists often suppose).

[361] *Id.* at 45 ("Of course there is plenty of influence and modelling and taking of hints among composers, apart from countless regurgitations.").

[362] *See id.* ("The present article is directed against, not the existence of [copying], but the insistence that *every* likeness means a connection."); *see also* Kyle Gann, *Things Composers Can Do*, POSTCLASSIC BLOG ARCHIVE (Jan. 14, 2015), https://www.artsjournal.com/postclassic/2015/01/things-composers-can-do.html [https://perma.cc/57EU-XXSE[] (complaining about the ease with which musicologists attribute borrowing to Ives without adequate proof).

whether that evidence indicated misuse or defect as the cause of a tire blowout.[363] Forensic musicologists examine musical evidence of potential copying (i.e., similarities) and opine whether those manifestations indicate copying or coincidence as the explanation.

As noted earlier, Carlson rendered a plausible opinion that was properly considered inadmissible because his analysis did not reflect a sufficiently reliable methodology or experience. Nothing indicated how principles of science, engineering standards, or accepted methods guided his analysis or observations. His testimony was therefore too dependent on personal fiat to pass the requirements of Federal Rule of Evidence 702.[364]

A forensic musicologist's opinion about copying is no different. If the musicologist testifies that a certain combination of similarities establishes copying, one must ask why these similarities are sufficient. Why not more or fewer similarities? Indeed, how does the expert even know that the type of similarity deemed sufficient even matters? Because musicology contains no method for answering these questions, the musicologist's opinion is as dependent on personal fiat as Carlson's. The factfinder is being asked to take the expert's word for a conclusion that musicology cannot reliably explain. This establishes her opinion as the type of insufficiently supported personal assertion excluded by Federal Rule of Evidence 702.[365]

Comparison with other similar forms of testimony reinforces this conclusion. For example, consider the evidentiary treatment of opinions by forensic document examiners about the source of handwriting, which also depend on an examiner's perception and judgment about similarity.[366] When compared to forensic musicology, forensic document examination comes with safeguards that forensic musicology

[363] *See supra* note 233 and accompanying text.

[364] *See supra* notes 238–240 and accompanying text.

[365] *See supra* note 31 (listing cases standing for the proposition that the Federal Rules of Evidence exclude expert testimony that rests too heavily on the expert's personal opinion or beliefs).

[366] *See, e.g.*, United States v. Foust, 989 F.3d 842, 846 (10th Cir. 2021) (comparing handwriting to fingerprint analysis and noting that both rely on the subjective judgment of the analyst); United States v. Crisp, 324 F.3d 261, 271 (4th Cir. 2003) (describing how handwriting expert examined similarities shared by known exemplars of defendant's handwriting and handwriting on a document).

opinions about copying do not. Forensic document examiners receive training and certification from institutions devoted to the discipline.[367] Forensic musicologists do not have access to such institutional training or credentialing. The Scientific Working Group for Forensic Document Examination and the American Association of Forensic Science publish standards designed to encourage reliability and accuracy.[368] Similar organizations and standards do not exist for forensic musicology.[369] And finally, handwriting analysis has been the subject of academic study for reliability[370] while forensic musicology has not.

[367] Among educational institutions, Eastern Tennessee State offers a certificate program: *Forensic Document Examination*, E. TENN. STATE UNIV., https://www.etsu.edu/online/graduate-certificates/forensic-document-examination.php (last visited July 24, 2024) [https://perma.cc/VUJ5-2ECA]; *see also, e.g.*, AM. ACAD. OF FORENSIC SCIS., https://www.aafs.org (last visited July 24, 2024) [https://perma.cc/XUS7-FFW6]; AM. BD. OF FORENSIC DOCUMENT EXAM'RS, https://abfde.org (last visited July 24, 2024) [https://perma.cc/YXK7-6E2G] (offering certification); BD. OF FORENSIC DOCUMENT EXAM'RS, https://www.bfde.org (last visited July 24, 2024) [https://perma.cc/893X-XAP9] ("[E]stablished to administer a professional certification program . . . ."); INT'L SCH. OF FORENSIC DOCUMENT EXAMINATION, https://internationalschool.us (last visited July 24, 2024) [https://perma.cc/39W4-6DRF] (offering remote education course of certification).

[368] AM. ACAD. OF FORENSIC SCIS., *supra* note 367; SCI. WORKING GRP. FOR FORENSIC DOCUMENT EXAMINATION, https://www.swgdoc.org (last visited July 24, 2024) [https://perma.cc/NH6N-QGPA]; *see also, e.g.*, ASTM INT'L, https://www.astm.org/ (last visited July 24, 2024) [https://perma.cc/6HQJ-ZQ5G]; *A Simplified Guide to Forensic Document Examination*, FORENSIC SCI. SIMPLIFIED, https://www.forensicsciencesimplified.org/docs/how.html (last visited July 24, 2024) [https://perma.cc/J9RQ-DVAM]; Mark Songer, *Forensic Document Examination*, ROBSON FORENSIC (Nov. 24, 2015), https://www.robsonforensic.com/articles/forensic-document-examination-expert-witness [https://perma.cc/DH5P-RD5NPERMA].

[369] The American Musicological Society lists forensic musicologists "as a public service." The listing's preface contains the admonition "let the user beware," presumably to let the reader know that the AMS neither credentials, trains, nor certifies the experts on the list. *See Forensic Musicology*, AM. MUSICOLOGICAL SOC'Y, https://www.amsmusicology.org/page/ForensicMusicology?&hhsearchterms=%22forensic+and+musicology%22 (last visited July 24, 2024) ; *see also* Fishman & Garcia, *supra* note 272, at 1204 ("[E]very expert with whom we spoke could think of no such networking outlet: no formal professional organization, no annual conference, no series of happy hours or meet-ups.").

[370] *See* R. Austin Hicklin, Linda Eisenhart, Nicole Richetelli, Meredith D. Miller, Peter Belcastro, Ted M. Burkes, Connie L. Parks, Michael A. Smith, JoAnn Buscaglia, Eugene M. Peters, Rebecca Schwartz Perlman, Jocelyn V. Abonamah & Brian A. Eckenrode,

The relatively developed nature of forensic document examination has convinced some courts to admit handwriting opinions under Federal Rule of Evidence 702.[371] However, other courts remain unconvinced about the reliability of such evidence and exclude it because handwriting analysis still depends too heavily on the subjective perception and judgment of the examiner. Their reticence mirrors concerns raised in the Advisory Committee Notes to Federal Rule of Evidence 702, which states, "Expert opinion testimony regarding the weight of feature comparison evidence must be limited to those inferences that can reasonably be drawn from a reliable application of the principles and methods."[372]

Disagreement over its admissibility shows that expert handwriting analysis straddles the line between admissible and inadmissible expert opinion under Federal Rule of Evidence 702. However, forensic musicology opinions about copying fall clearly on the inadmissible side of that line because forensic musicology has none of the safeguards that support admissibility of handwriting opinions.[373]

Remember, courts that admit handwriting opinions do so because training, promulgation of standards, and experimental testing offer reassurance that those opinions flow from expert knowledge instead of personal fiat.[374] Forensic musicology, however, lacks those reassurances and is therefore even more dependent on the subjective judgment and perception of the examiner.[375] One might be tempted to overlook

Accuracy and Reliability of Forensic Handwriting Comparisons, 119 NO. 32 PROC. NAT. ACAD. SCI. U.S. (2022) (studying accuracy and reliability of handwriting comparison by forensic document examiners).

[371] *See* United States v. Mallory, 902 F.3d 584, 595 (6th Cir. 2018) (affirming district court decision to admit handwriting expert); *supra* note 366.

[372] FED. R. EVID. 702 Advisory Committee's Notes. For cases excluding opinions about handwriting, see Crew Tile Distrib., Inc. v. L.A., Inc., 763 F. App'x 787, 790 (10th Cir. 2019) (excluding testimony of handwriting expert); Almeciga v. Ctr. Investigative Reporting, Inc., 185 F. Supp. 3d 401, 407-08, 418 (S.D.N.Y. 2016) (excluding testimony of handwriting expert even though expert relied upon a commonly used method to render opinion); United States v. Johnsted, 30 F. Supp. 3d 814, 816 (W.D. Wis. 2013) (holding handwriting analysis insufficiently reliable under Federal Rules of Evidence 702 and 403).

[373] *See* Fishman & Garcia, *supra* note 272, at 1204 (noting lack of any professional organization or other formal networking for forensic musicologists).

[374] *See supra* notes 366, 371.

[375] *See supra* notes 369–370 and accompanying text.

training, standardization, and testing as less important than assurances provided by forensic musicologists (or other experts) that their opinions are valid and reliable. We should not forget, however, that experts testifying about forensic matters routinely provide such assurances only to have their testimony discredited at a later date.[376] It is therefore difficult to see why courts should allow musicologists to opine about the existence of copying under Federal Rule of Evidence 702.

III. THE PROPER USE OF FORENSIC MUSICOLOGY

Courts misuse forensic musicology in copyright cases because they make two errors. First, they adopt confusing and inaccurate elaborations of copyright law that collapse the distinction between proof of copying and proof of improper appropriation.[377] This results in the use of forensic musicology opinions that should be inadmissible because they are not helpful to the jury and prejudicially confusing under Federal Rule of Evidence 403.[378] Second, they do not fully appreciate the poor foundation that musicology provides for the reliable identification of copying.[379] This results in the admission of opinions that do not meet the requirements of Federal Rule of Evidence 702.[380]

The proper use of forensic musicology begins with getting the law right. Courts should abandon unhelpful and confusing elaborations of *Arnstein* that collapse the distinction between copying and improper appropriation. Among other things, courts should stop using "substantial similarity" to denote the level of proof required for both copying and improper appropriation. Terms like "sufficient probative similarity" and "sufficient aesthetic similarity" would clearly signal to courts and juries alike that the evidence and inferences required for copying are distinct from those required for improper appropriation. Terminology like this should also mean the end of the Ninth Circuit's

[376] The Innocence Project reports that "misapplied forensic science" contributed to nearly a quarter of all wrongful conviction cases since 1989. *Misapplication of Forensic Science*, INNOCENCE PROJECT, https://innocenceproject.org/misapplication-of-forensic-science/ (last visited July 24, 2024) [https://perma.cc/M5EG-S3MK].

[377] *See supra* Part II.B.

[378] *See supra* notes 276–281 and accompanying text.

[379] *See supra* notes 283–361 and accompanying text.

[380] *See supra* notes 364–376 and accompanying text.

extrinsic and intrinsic analysis for infringement. As shown earlier, that analysis is widely criticized and also collapses the distinction between copying and improper appropriation.[381] Its end would be a salutary development for copyright law and help courts use forensic musicology correctly.

Once courts have the law straight, they must make sure that forensic musicologists testify about matters relevant only to copying, not improper appropriation. Courts should therefore generally allow forensic musicologists to identify formal musical similarities and offer information properly relevant to evaluating those similarities as evidence of copying. Experts should not opine that two works make the same overall aesthetic impression, nor should they testify that two works are "substantially similar." They should instead testify about the musical characteristics of individual similarities (such as rarity or aesthetic significance) that affect inferences about copying.

Policing this limit may not prove as easy as it sounds because similarities may be simultaneously probative and aesthetic.[382] An expert may want to testify that the similar overall aesthetic impression of two works is probative of copying. Such proposed testimony would technically address probative similarity, but it would also prejudicially suggest that aesthetic similarity establishes copying or affects the jury's assessment of improper appropriation. Courts could deal with this problem in two different ways.

First, courts could simply prohibit this testimony. Proponents of the testimony might object that this excludes evidence about copying, but closer examination reveals workarounds that avoid testimony about aesthetics. An expert claiming that two works are aesthetically similar should be able to explain the reasons for this in formal musical terms. For example, the works could be aesthetically similar because they share the same melody, rhythm, and tempo. Once the expert identifies these formal similarities, she can analyze whether they are probative of copying without mentioning their overall aesthetic effect. This would improve the probative value of the expert's opinion while greatly reducing the prejudicial dangers associated with discussing aesthetics.

[381] *See supra* notes 64, 90.

[382] *See supra* notes 51–59 and accompanying text (noting potential overlap between aesthetic and probative similarity).

This would probably allow the expert's opinion to avoid exclusion for reasons of prejudice and confusion, but problems of reliability would of course remain.

Second, courts could permit this testimony but bifurcate the trial so that the issue of copying is tried first.[383] This will focus the jury's attention entirely on copying, making it far less likely that the expert's opinion will influence improper appropriation. It will also have the benefit of making sure that courts themselves get the law right. Courts would try the improper appropriation issue only after a jury has found that copying exists. Granted, there will remain some risk that the jury will determine copying on the basis of aesthetics, but this degree of prejudice might be acceptable, particularly with an appropriate limiting instruction.

Once courts properly confine forensic musicology testimony to the issue of copying, attention should turn to the challenges of reliability. Here, courts should not allow experts to opine that copying does or does not exist. Such testimony lacks adequate methodological support and depends too heavily on the expert's personal fiat. Instead, courts should restrict experts to testifying about the relative or comparative strength of inferences about copying drawn from similar features *when an appropriate foundation exists*. So, for example, an expert (probably a defendant's expert) should be allowed to testify that inferences of copying are relatively weak because a particular musical similarity at issue commonly appears in other works if the expert can provide examples of the feature to support the opinion.[384] Alternatively, an expert (probably a plaintiff's expert) might try to claim that the inference of similarity is relatively strong because the musical feature in question is rare, but such testimony remains problematic, at least for

[383] Alternate possibilities include special jury forms or instructions about the proper understanding to be given to expert testimony. I think that, while plausible, neither will be as effective as bifurcation for directing the jury's attention to copying. Proper bifurcation would ensure that juries think only about copying in the absence of improper appropriation, while jury forms and instructions try to guide juries after hearing about both. Thus, only bifurcation completely avoids confusion about evidence associated with improper appropriation.

[384] *See* Fishman & Garcia, *supra* note 272, at 1197 (stating the "unextraordinary" proposition that examples of other works containing the similarities complained of by the plaintiff renders coincidence a more likely explanation).

the time being. An expert testifying about rarity on the basis of his listening and study should be challenged for reasons of comprehensiveness or representativeness, and courts should exclude or significantly limit such testimony unless the proponent can lay a foundation that the expert has reviewed an appropriately large and representative subset of the relevant corpus.[385] It bears repeating, however, that even if the foundation is adequate, no expert should testify that copying did or did not occur.[386]

[385] *See supra* notes 268–272 and accompanying text. Going forward, the emergence of searchable music databases may make at least some testimony of this sort more reliable. Such databases could, at least in theory, become sufficiently comprehensive that computer-assisted searches provide reliable evidence about certain types of musical features. For example, the Josquin Research Project hosts a growing searchable online database that contains a growing collection of complete Renaissance music scores. It allows the user to "identify every instance of a given melodic and/or rhythmic pattern." *About the Josquin Research Project*, JOSQUIN RSCH. PROJECT, https://josquin.stanford.edu/about/ (last visited July 24, 2024) [https://perma.cc/E3FM-PKEZ]. Over time, it is reasonable to imagine that similar databases will exist for other genres of work. At this time, IMSLP (a music database that devoted to "sharing the world's public domain music") permits searching for scores by melody. INT'L MUSIC SCORE LIBR. PROJECT, https://imslp.org/wiki/Main_Page (last visited July 24, 2024) [https://perma.cc/N4T6-2D5B]. Ironically, a major obstacle to the creation of such databases could be copyright law itself because copyright holders would have plausible claims of copyright infringement based on the reproduction of music in the databases. *See* 17 U.S.C. § 106(1) (reserving to copyright holder the exclusive right to reproduce the copyrighted work in copies or phonorecords).

[386] Litigants will surely offer expert testimony that functions as an opinion about the existence of copying, even if the expert does not directly so state. Judges should consider such testimony carefully to ensure that experts do not exceed the boundaries of reliability or confuse the jury. It bears noting that at least one professional association has voluntarily limited the types of opinions its members should render in a somewhat analogous situation. Forensic odontologists are sometimes asked to opine whether a particular individual's teeth are responsible for a bitemark. One might imagine that such experts could render opinions ranging from exclusion of the individual's teeth as the source of the bitemark to positive identification of the individual's teeth as the source. However, the American Board of Forensic Odontology recognizes that opinions excluding a potential source of a bitemark are much more reliable than those identifying a source. Accordingly, American Board of Forensic Odontology guidelines permit only three opinions: (1) source excluded as having made the bitemark, (2) source not excluded, and (3) inconclusive. In other words, an expert adhering to these guidelines should not opine that the potential source left the bitemark. *See* AM. BD. OF FORENSIC ODONTOLOGY, ABFO BITEMARK METHODOLOGY STANDARDS AND GUIDELINES (2016),

If courts adopt these suggested changes, the litigation of music copyright infringement cases will improve. Clarity about the law and bifurcation of the trial will ensure that juries properly consider all elements of the case. Additionally, limiting forensic musicologists to the identification of probative similarity and properly supported opinions about the rarity of relevant similarities will greatly lower the risk of confusing juries. Parties to copyright litigation will no longer be able to offer unreliable expert testimony about inferences that should be reserved for summation. This will help juries understand the evidence they do hear and the inferences they must make in order to find copyright infringement.

IV. POSSIBLE OBJECTIONS

There are, of course, plausible objections to the recommendations made here. First, one might claim that forensic musicologists have exceeded the field of musicology, creating analytical methods that permit reliable opinions about copying. Second, one might take the position that there is little harm in allowing experts to opine that copying does or does not exist because each side can present its own expert and cross-examine the other's. Third, one could object that the proposals made here risk improperly tipping the scales in music infringement cases towards defendants by depriving plaintiffs of testimony crucial to carrying their burden of proof on copying. Fourth, some may object that without expert conclusions about copying, juries will engage in uninformed speculation that could be avoided with more comprehensive expert opinions. This Part will consider each of these possibilities.

http://abfo.org/wp-content/uploads/2016/03/ABFO-Bitemark-Standards-03162016.pdf [https://perma.cc/PZY2-W6Q9]; AM. BD. OF FORENSIC ODONTOLOGY, STANDARDS AND GUIDELINES FOR EVALUATING BITEMARKS 17 (2018), http://abfo.org/wp-content/uploads/2012/08/ABFO-Standards-Guidelines-for-Evaluating-Bitemarks-Feb-2018.pdf [https://perma.cc/68EX-JPLM].

A. Has Forensic Musicology Exceeded Musicology by Developing Reliable Methods of Its Own?

Forensic musicologists sometimes assert that their analyses follow "standard musicological procedure,"[387] and one might claim that the existence of these procedures establishes reliability because *Daubert* stated that general acceptance within a field serves as an important factor in determining reliability.[388] This argument has facial appeal, but it fails to support the admissibility of expert opinions about copying. As an initial matter, references to standard procedure often involve well-accepted musicological techniques that help identify formal similarities between two works such as musical transcription or transposition.[389] As such, these accepted techniques support opinions about basic information that have already been identified as admissible.[390] However, the admissibility of testimony about formal similarities does not make inferences about copying admissible. To the extent that experts claim a standard musicological procedure for making reliable inferences about copying, the claimed standard procedures are still not sufficiently reliable to pass Federal Rule of Evidence 702.

For example, consider the thoughtful and perceptive proposal made in 2023 by Professor Joe Bennett in *We Can Work It Out: Methods in Forensic Musicology*.[391] Bennett proposes that a correctly performed

[387] *See* Declaration of Dr. Alexander Stewart in Support of Opposition to Defendants' Motion for Summary Judgment, Exhibit 15 at 2-7, Smith v. Tesfaye (No. 2-19-cv-02507-PA-MRW) (2020) (expert witness stating that he follows "standard musicological procedure"); Anthony Ricigliano, A Musicological Report Regarding Baghon Mein Bahar Nai by Mohd Rafi, Lata Mangeshkar and Put You On The Game by Mosely, Taylor and Hills, Roberts v. Gordy (No. 1:13-cv-24700-KMW), Document 336-2 at 2 (2015) (same). For a discussion of "standard musicological procedure," see LEO, *supra* note 5, at 151-54.

[388] Daubert v. Merrell Dow Pharms. Inc., 509 U.S. 579, 594 (1993).

[389] *See* LEO, *supra* note 5, at 152 (describing standard musicological practice as "common, often notation-based analytical techniques"); Ricigliano, *supra* note 387, at 2 (referring to standard musicological procedure for purposes of comparing formal similarities between works); Stewart, *supra* note 387, at 3 (referring to standard musicological procedure in the context of transcription and transposition).

[390] *See supra* Part II.C.1 (identifying as admissible expert testimony about formal similarities between two works).

[391] Joe Bennett, *We Can Work It Out: Methods in Forensic Musicology*, *in* MUSIC BORROWING AND COPYRIGHT LAW: A GENRE-BY-GENRE ANALYSIS 57 (Enrico Bonadio & Chen Wei Zhu eds., 2023).

analysis of copying should contain three parts. First, the expert should identify similarities and make a judgment about whether those similarities are probative of copying.[392] This parallels the quasi-probabilistic inference described by Burkholder.[393] Second, the expert should identify and exclude from consideration similarities arising from "commonplace elements."[394] This corresponds to the notion that similarities arising from well-known musical conventions constitute poor evidence of copying. Third, the expert should identify and exclude similarities that also exist in other music on the theory that such existence strengthens the explanation of independent coincidental creation.[395] Bennett argues that this method avoids many errors that forensic musicologists presently make.[396] Although Bennett does not consider how his proposal affects the admissibility of forensic musicology opinions, one could offer his method as one that leads to reliable conclusions about copying.

Professor Bennett's proposal resembles Professor Burkholder's work in that both describe good musicological analyses of suspicious similarities. However, because of the limitations inherent in musicology, Bennett's method fares no better than Burkholder's when it comes to reliability required by the Federal Rules of Evidence. Indeed, his proposal illustrates the value of distinctions made by this Article about admissible expert testimony, inadmissible expert speculation, and inferences that should be reserved for the jury.

Note that the three steps identified by Bennett rest on information already identified as admissible, namely the identification of similarities and testimony about their prior existence. Such evidence is admissible as long as the proper foundation gets laid. However, to the extent that Bennett contemplates experts opining that this information does or does not establish copying, familiar concerns about reliability arise

[392] *See id.* at 66-67.

[393] *See id.* at 67 (using probabilistic intuition); Burkholder, *Musical Borrowing*, *supra* note 7, at 227-28, 256-57 (describing the relationship between probability and likelihood of copying).

[394] Bennett, *supra* note 391, at 67.

[395] *Id.* at 68-69.

[396] *Id.* at 69-71 (arguing that "prior art research is . . . the most robust methodology for rebutting spurious accusations of infringement").

when the expert decides whether the provided information makes copying more likely than not. Which similarities matter? Which similar features in other works matter? How many similarities are enough? As noted earlier, musicology does not give musicologists answers to these questions, so personal fiat becomes determinative.

These observations show why the correct evidentiary approach allows experts to provide the information contemplated by Bennett while excluding opinions about the implications of that evidence. As noted earlier, the provision of the information contemplated by Bennett rests on a sufficient foundation to meet Rule 702's requirement of reliability. However, such a foundation does not exist when musicologists tell juries what to infer from that information. Determinations of that sort should be made by the jury after argument from counsel at summation. Indeed, lay jurors are fully capable of making the sorts of inferences contemplated by Bennett. They understand that the prior existence of identified similarities in other works weakens any claim of copying.

B. What Is the Harm of Having Forensic Musicologists Give Opinions About Copying?

One could object to the proposals made here on the ground that our adversarial process ameliorates any unreliability associated with opinions about copying. By the time juries hear an expert identify or dispute the existence of similarities, provide information about rarity, and explain the musical significance of similarities, they will know what the expert thinks. Allowing expert opinions about copying will therefore have little effect on jury deliberation.[397] Moreover, even if these opinions are methodologically questionable, they are unlikely to mislead juries because each side will present its own expert and cross-examine the other side's musicologist. This will draw attention to any errors either expert makes, thereby giving juries an accurate picture of the weight to which forensic musicology opinions are entitled.

This argument is plausible but mistaken. As an initial matter, if these opinions really did not affect juries, litigants would not work so hard to

[397] *See* Lee & Moshirnia, *supra* note 5, at 735-83 (presenting results of empirical testing about the effect of expert testimony in music copyright cases suggesting limited effect of such testimony).

present them. Furthermore, there is a big difference between presenting unreliable opposing experts to juries and giving juries the benefit of experts with divergent but reliable opinions.

Under current practice, the jury hears opposing musicologists who assert that their expertise allows them to tell whether the defendant copied from the plaintiff's work.[398] The plaintiff's expert states with certainty that copying has occurred, while the defendant's expresses equal certainty that it has not. Crucially, neither party will inform the jury that musical analysis does not function as the experts advertise. Indeed, each party has a strong incentive to confidently assert that its expert is right, and this means accepting the premise that musicologists can accurately and reliably identify when copying has occurred. The jury will never find out that reliable conclusions about copying exceed the epistemological limits of musicology and that, at best, musicology provides terminology and basic information on which the expert can base personal hunches.

This seriously misleads the jury. Jurors are expected to believe one expert over the other because that expert convinces the jurors that her analysis is musicologically correct and therefore reliable when musicological conclusions about copying actually depend on personal fiat.[399] This implies sending jurors on a fool's errand by asking them to choose between two foundationally unreliable opinions. The result would be similar (though not identical) to admitting expert testimony by astrologers because cross-examination helps juries decide which astrologer had performed the more convincing astrological analysis.[400]

Limiting forensic musicology testimony as suggested here eliminates considerably the fool's errand. When nobody tells juries that a musicologist can give them the right answer, juries properly understand that the case in front of them is ambiguous and that neither party has iron-clad proof. Importantly, juries will stop overvaluing musicological testimony and focus instead on all of the available evidence, including the testimony and cross-examination of the parties. Verdicts rendered

[398] *See supra* Part II.B.

[399] *See supra* Part III.C.4.

[400] *See* Kumho Tire Co. Ltd. v. Carmichael, 526 U.S. 137, 151 (1999) (noting that applying generally accepted principles of astrology does not make an astrologer's testimony reliable).

this way may or may not be accurate, for there is no way to reliably tell when copying has occurred. However, the verdicts will not be fundamentally confused about the probative value of evidence presented.

C. Does Excluding Opinions About Copying Unfairly Favor Defendants?

Some, especially plaintiffs, may object that excluding expert opinions about copying will unfairly favor defendants. If juries understand that musicological proof of copying is highly speculative or they never hear musicological opinions about the existence of copying, they may become more likely to conclude that the plaintiff cannot carry its burden of proof. This is an interesting concern in that its underlying observation may be correct. If music copyright plaintiffs rely heavily on expert witnesses to carry the burden of proof on copying, limiting this testimony would presumably weaken the plaintiff's case in chief, perhaps to the point that summary judgment or directed verdicts in favor of defendants become possible. However, even if the underlying observation is true, unfairness does not necessarily follow.

It is fair to impose a basic burden of proof on the plaintiff because doing so constitutes a basic feature of our litigation system. If the plaintiff has only equivocal evidence to support his position, a judge or jury may conclude that the plaintiff has a poor case. That is precisely what should happen. Granted, one could arguably characterize a trial as "fair" if both sides can present any evidence they desire, regardless of its probative value, prejudice, or reliability. However, the Federal Rules of Evidence do not take this position and limit both sides to evidence that is relevant and not unduly prejudicial or (in the case of expert testimony) unreliable. Accordingly, if certain types of forensic musicology testimony are unduly prejudicial or unreliable, there is nothing unfair about eliminating them from trials. If any disproportionate effect on plaintiffs exists, that is because current practice has given plaintiffs the disproportionate advantage of meeting burdens of proof with inadmissible evidence. Forcing plaintiffs to go without this evidence simply restores the sort of fair trial contemplated by the Federal Rules of Evidence.

Additionally, it is highly unlikely that the proposal made here will cause plaintiffs to lose cases at summary judgment improperly. A court

should grant summary judgment for the defendant only when the plaintiff's evidence is so weak that a jury cannot reasonably find in the plaintiff's favor.[401] The proposal made here leaves plaintiffs plenty of leeway to avoid losing at summary judgment or directed verdict. Plaintiffs can use expert witnesses to identify similarities between works and to provide information that supports an intuitive inference of copying. Unless this evidence is very weak, it is likely that judges will deny summary judgment or directed verdict because juries can rationally infer copying from this evidence.[402] Of course, if the evidence is so weak that the judge, even after expert testimony, cannot see a rational decision in the plaintiff's favor, then summary judgment or directed verdict would be justified.

D. Are Expert Opinions About Copying Necessary to Prevent Speculation by the Jury?

Finally, it is possible to argue that expert conclusions about copying are helpful to juries because they prevent speculation. Again, I respectfully disagree. As an initial matter, forensic musicologists engage in considerable speculation themselves. It is doubtful that any forensic musicologist is more thoughtful and careful (or as accomplished) as Professor Burkholder, and even Professor Burkholder cannot articulate a method for reliably determining when suspicious similarity becomes more likely than not.[403] Musicologists may engage in more sophisticated or educated speculation than lay jurors, but it is still unreliable speculation that should not be used as evidence.

Furthermore, when it comes to unreliable speculation, our system of litigation holds expert witnesses to a higher standard than juries. By

[401] *See* FED. R. CIV. P. 56(a) (authorizing summary judgment when there is no genuine dispute about a material fact and moving party is entitled to judgment as a matter of law); Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 248 (1986) (stating that summary judgment is not appropriate if a reasonable jury could find for the non-moving party).

[402] In this regard, it is instructive to note that *Arnstein v. Porter* itself reversed a district court's grant of summary judgment in the defendant's favor. The plaintiff's case was weak, but the Second Circuit considered the similarities enough to warrant going to a jury. *See supra* notes 36–42 and accompanying text.

[403] *See supra* note 333 and accompanying text (analyzing how Burkholder's method contains no guidance about when the items he recommends analyzing indicate copying).

imposing requirements of expert qualification and reliability,[404] the Federal Rules of Evidence actively manage the submission and content of expert testimony to prevent unreliable, speculative opinions from affecting trials. By contrast, nothing beyond the rules regarding judgments notwithstanding the verdict prevent a jury from deliberating as it sees fit.[405] Indeed, courts generally protect jury deliberations as secret, making it extremely difficult (if not impossible) to manage how jurors decide cases.[406]

CONCLUSION

The stakes are high when one composer sues another for copyright infringement. The plaintiff impugns the defendant's artistic integrity by alleging that the defendant claims to have written music actually created by the plaintiff. Should the plaintiff succeed in his suit, his recovery can include monetary damages as well as injunctions against future sales, performance, and licensing of the defendant's work. Courts may even order the impounding and destruction of copies of the defendant's work.[407] These consequences may well be appropriate if the defendant did indeed commit infringement. However, if courts mistakenly declare the defendant an infringer, the consequences are horribly unjust because they amount to the defendant's loss of her work and — by extension — her livelihood. Because of this, our legal system should require solid evidence of infringement before imposing these consequences on a defendant who denies infringing.

Under current practice, courts rely heavily on forensic musicology testimony to determine if infringement has occurred. However, as a matter of evidence law, the field of musicology provides a rather shaky

[404] *See* FED. R. EVID. 702 (imposing requirements of reliability and expert qualification).

[405] *See* FED. R. CIV. P. 50 (giving court the power to override jury verdict only if a reasonable jury would not have a legally sufficient evidentiary basis for its verdict).

[406] *See* FED. R. EVID. 606(b) (prohibiting jurors from testifying about any juror's "mental processes" during an inquiry about the validity of a verdict); United States v. Thomas, 116 F.3d 606, 618-20 (1997) (explaining that the secrecy of jury deliberation is important); United States v. Olano, 507 U.S. 725, 737 (1993) (noting the "cardinal principle" that jury deliberations must be secret).

[407] *See* 17 U.S.C. §§ 502–05 (establishing remedies for copyright infringement).

foundation for considerable amounts of forensic musicology testimony presently accepted by courts. Courts should therefore meaningfully limit forensic musicology testimony to reliable information and opinions that focus the jury's attention on legally relevant issues.

This recommendation is not a general attack against musicology. I am only analyzing the legal evidentiary consequences of inherent limitations that all musicologists face when studying copying. There is nothing wrong with musicologists doing the best they can with the tools they have. This does not mean, however, that those tools should be used to determine legal rights, especially when the consequences of error are significant. It is one thing for those studying music to debate the significance of musical similarities that raise the possibility of copying. Such dialogue enriches our understanding of music, and any conclusions remain open to revision as additional evidence or understandings emerge. The adjudication of legal rights, however, is final. For that reason, our judicial system properly requires that expert testimony be sufficiently reliable before courts use it to make decisions that cannot be revised. That is why it is worth scrutinizing the limits of forensic musicology to make sure that courts properly adjudicate music copyright infringement cases.