THE JOHN MARSHALL
REVIEW OF INTELLECTUAL PROPERTY LAW

# BLURRING THE LINE: AN EXAMINATION OF TECHNOLOGICAL FACT-FINDING IN MUSIC COPYRIGHT LAW

JEREMY AREGOOD

ABSTRACT

The result of Williams v. Bridgeport Music, Inc. highlights a major issue in musical plagiarism fact-finding. Different circuits employ different tests for fact-finding, however all the tests involve some form of "objective" criteria that is guided by expert witnesses who perform musical analyses. Because expert witnesses influence their analysis with their own subjective interpretations of the music, and because juries are not fully aware of the distinction between objective and subjective analysis, juries have a distinct possibility of returning a verdict that contradicts the evidence and public policy. New advancements in technology and computation may assist courts in evaluating the objective similarity factors of musical plagiarism. This comment examines the music analytic technology being developed today, explores the potential applications as well as the implications of adopting new technology to improve music copyright law, and provides an understanding of the increasingly blurred line between objective and subjective music copyright litigation.



*Cite as* Jeremy Aregood, *Blurring the Line: an Examination of Technological Fact-Finding in Music Copyright Law*, 16 J. MARSHALL REV. INTELL. PROP. L. 115 (2016).

# BLURRING THE LINE: AN EXAMINATION OF TECHNOLOGICAL FACT-FINDING IN MUSIC COPYRIGHT LAW

JEREMY AREGOOD

# BLURRING THE LINE: AN EXAMINATION OF TECHNOLOGICAL FACT-FINDING IN MUSIC COPYRIGHT LAW

JEREMY AREGOOD*

## I. INTRODUCTION

The songwriting duo of Pharrell Williams and Robin Thicke is no stranger to controversy. In 2013, the internet was abuzz with articles decrying the sexually objectifying lyrics contained in their newest hit. Artists often welcome this sort of free publicity, but another controversy was brewing that threatened to jeopardize their entire joint project. Responding to accusations of plagiarism, Williams and Thicke filed an action for declaratory relief[1] against the Gaye family, hoping to put to rest the rumors that the record breaking[2] single, "Blurred Lines," unlawfully borrowed from Marvin Gaye's "Got to Give it Up." The Gaye family filed counterclaims for copyright infringement and aimed to dismiss the motion for declaratory relief.[3] The Central District of California acknowledged that the Gaye family had provided sufficient evidence that the musical phrases in question were not merely "scènes à faire."[4]

Having established that the Gaye family provided enough evidence to contest a triable issue of material fact, the court empaneled a jury as finder of fact and a trial began.[5] The district court instructed the jury to decide whether there was a "substantial similarity"[6] between the musical excerpts in contest by employing an

---

* © Jeremy Aregood 2016. Author Bio. J.D Anticipated May 2017, The John Marshall Law School; B.A. Economics, Northwestern University. Thank you to the angelically patient and hardworking RIPL staff, who helped keep me on task throughout this experience. Special thanks to my mother, Rhonda Stuart, for encouraging me to join RIPL and for listening to me talk at length about my topic, even in the smallest hours of the morning.

[1] Williams v. Bridgeport Music, Inc., 2015 U.S. Dist. LEXIS 97262, at *2 (C.D. Cal. July 14, 2015).

[2] "Blurred Lines" remained at the top of Billboard's Hot R&B/Hip Hop Songs chart for sixteen weeks—the longest amount of time since Nielsen began collecting data for R&B in 1942. *See* Rauly Ramirez, *Robin Thicke's 'Blurred Lines' Breaks Record Atop Hot R&B/Hip-Hop Songs*, BILLBOARD (Sept. 25, 2013), http://www.billboard.com/articles/columns/the-juice/5733206/robin-thickes-blurred-lines-breaks-record-atop-hot-rbhip-hop. "Blurred Lines" also sold the most single-track digital downloads of 2013 at 14.8 million units. *See* Digital Music Report 2014, International Federation of the Phonographic Industry, http://www.ifpi.org/downloads/Digital-Music-Report-2014.pdf. (last visited Dec. 21, 2015).

[3] *Williams*, 2015 U.S. Dist. LEXIS 97262, at *2.

[4] *See* Swirsky v. Carey, 376 F.3d 841, 846 n.11 (9th Cir. 2004) (stating that "scènes à faire are common expressions indispensable to the expression of particular ideas in a relevant field; they are treated as unprotectable by copyright, in the manner of ideas"). The scènes à faire doctrine stems from the limited nature of certain forms of artistic expression in a medium, such as the famous twelve-bar blues progression. *See, e.g.*, Garth Alper, *How the Flexibility of the Twelve-Bar Blues Has Helped Shape the Jazz Language*, 45 C. MUSIC SYMP. 1 (2005); *see also* Williams v. Bridgeport Music, Inc., 2014 U.S. Dist. LEXIS 182240, at *14 (C.D. Cal. Oct. 30, 2014) (plaintiffs did "not contest Defendants' ownership of valid copyrights in 'Got to Give It Up'").

[5] *Williams,* 2015 U.S. Dist. LEXIS 97262, at *3.

[6] *See id.* at 64 (stating 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"). Because it is often prohibitively difficult to prove direct copying of a work, many jurisdictions require that a plaintiff prove only that defendant had access to the protected work

"extrinsic" test[7] and "intrinsic" test.[8] The jury relied heavily on the extrinsic test,[9] which measures "external, objective criteria," and involves "analytical dissection of a work and expert testimony."[10] After seven days of trial and two days of deliberation the jury returned a verdict, finding "by a preponderance of the evidence that the Thicke Parties 'Infringed the Gaye Parties' copyright in the musical composition 'Got to Give it Up' in 'Blurred Lines,'"[11] and awarded a total of $7,378,647.19 in actual damages and profits received to the Gayes.[12]

A number of attorneys have questioned the jury's finding of copyright infringement, calling into doubt the instructions and evidence presented to the jury.[13] On appeal of the judgment, the Thicke Parties questioned the evidence presented by expert witnesses, which satisfied the extrinsic test, such as the opinion that "certain elements" would be "understood by a musician to be present" in the copyrighted deposit copy of 'Got to Give it Up,'[14] or that the use of certain lyrics suggested a musical theme common to both works.[15] Despite the efforts to quantify and objectively analyze song similarities, music remains a fundamentally subjective and phenomenological

---

and that the works are substantially similar. *Id*. *See also* Whelan Assocs. v. Jaslow Dental Lab., Inc., 797 F.2d 1222, 1231-32 (3d Cir. 1986) (declaring that "as it is rarely possible to prove copying through direct evidence, copying may be proved inferentially by showing that the defendant had access to the allegedly infringed copyrighted work and that the allegedly infringing work is substantially similar to the copyrighted work").

[7] *Williams,* 2015 U.S. Dist. LEXIS 97262, at *65.

[8] *See id.*

[9] *See id.* at *68–69 (in fact, "[t]he Thicke Parties claim[ed] the jury never even conducted the intrinsic test, because it was not instructed on this test until after the close of evidence, and the jury never asked to hear any music during deliberations despite being told it could upon request"); *See infra* Part II for discussion of intrinsic tests.

[10] *Williams*, 2015 U.S. Dist. LEXIS 97262, at *65.

[11] *See id.*

[12] *See id.* at *4.

[13] *See* Harley Brown, *'Blurred Lines' Verdict: Music Lawyers Weigh In*, BILLBOARD (Mar. 11, 2015), http://www.billboard.com/articles/news/6495167/blurred-lines-verdict-music-lawyers-react (Thicke's attorney Dina Lapolt expressed concern that the jury may have misinterpreted being influenced by a song as having substantially copied it, and that it was hazardous to enforce copyright laws on a "genre or a style or a groove;" attorney Lawrence Iser remarked that the process of music production involves "borrow[ing] from the past to move the art forward"); *cf.* J. Peter Burkholder, *The Uses of Existing Music: Music Borrowing as a Field*, 50 NOTES 851, 870 (1994) (what appears to be specific borrowing may "represent an extreme case of a more widely shared procedure or tendency to use existing music"); *See also* Charles Cronin, *Comment for Pharrell Williams, et al. v. Bridgeport Music, et al.*, MUSIC COPYRIGHT INFRINGEMENT RESOURCE, http://mcir.usc.edu/inplay/Pages/williams.html (last visited Jan. 18, 2016) (claiming that "there is no legal or factual justification for the verdict," and that the jurors were swayed by a portrayal of the defendants as ne'er-do-well drug addicts and a portrayal of the plaintiffs as aggrieved saintly widows.)

[14] *Williams*, 2015 U.S. Dist. LEXIS 97262, at *18.

[15] The Gaye's musicologist offered the expert opinion that the words 'Up,' 'Down,' 'Shake,' and 'Round,' were lyrics common to both songs that exemplified "word painting:" the act of relating a musical theme or idea to words present in the song. *Id*. at *27–29. The Thicke parties also complained that because the comparison between the works relied on the sound recordings of the copies and not the copyright deposits which outline the musical scores, the comparisons were invalid and included non-copyrighted elements of the works. *Id*. at *15–16. The Thicke parties additionally demanded a hearing before and during trial to qualify the Gaye parties' expert musicologist Judith Finell, whose analysis "had no 'reliable basis in musicological practice,'" according to the Thicke parties. *Id.* at *17. This request was denied both times. *Id*.

experience, resulting in potentially confusing and prejudicial instructions for juries.[16] In light of the inherent impossibility of perfectly reconciling the subjective nature of music with the goal of objectivity in the administration of justice, the question remains: how might we shorten the gap between scientifically quantifiable audio and philosophically intangible music in the realm of copyright infringement? This article will do three things: (1) examine the music analytic technology being developed today; (2) explore the potential applications as well as the implications of adopting new technologies in order to improve music copyright law; (#3) provide an understanding of the increasingly blurred line between objective and subjective music copyright litigation.

## II. BACKGROUND

### *A. The Origin of Modern Music Copyright Law*

Modern copyright law has origins dating back to the invention of the printing press and the Statute of Anne—an English policy effectively granting copyright protection to "authors or proprietors" of "printed books and writings" from unconsented reproduction.[17] Technological advancement has always played a crucial role in changing the legislative attributes of an infringement claim, and the law often struggles to keep up with rapidly changing technological forces, especially in the music industry.[18]

The purpose of copyright law in music is to promote the advancement of the arts by incentivizing the creation of new and original musical works through various laws of protection, thereby fostering both creativity and economic prosperity.[19] The United

---

[16] *See* Durand R. Begault, *Analysis Criteria for Forensic Musicology,* 19 PROC. MEETINGS ON ACOUSTICS 1, 2 (2013) (explaining that there are few "standards, papers and methodologies" in musicology and that musical analysis presents problems that are difficult to "objectively analyze").

[17] *The Statute of Anne*, 8 ANN. C. 19 (1709); *see also The Copyright Act of 1790*, available at http://copyright.gov/history/1790act.pdf (the United States' first act relating to copyright law borrowed heavily from England's Statute of Anne, and is copied nearly verbatim; like the Statute of Anne it provides protection for printers).

[18] *See* Ben Depoorter, *Technology and Uncertainty: The Shaping Effect on Copyright Law*, 157 U. PA. L. REV. 1831 (2009) (concluding that "uncertainty over the application of existing copyright law to newly emerging technology catalyzes the actions of copyright owners and users," resulting in legal delay). *See also* Olufunmilayo B. Arewa, *From J.C. Bach to Hip Hop: Musical Borrowing, Copyright, and Cultural Context*, 84 N.C.L. REV. 547, 554 (2006) (noting that "[b]road technological changes since the late twentieth century have also profoundly influenced many areas, including music and copyright"). Olufunmilayo B. Arewa, *Recalibrating Copyright: Continuity, Contemporary Culture, and Change: Institute* for *Intellectual Property & Information Law Symposium: A Musical Work is a Set of Instructions,* 52 HOUS. L. REV. 467, 470 (2014) (contending that technological advancement and changing music production "increasingly challenge . . . notation in music copyright").

[19] *See* Michael Carroll, *Whose Music is it Anyway?: How We Came to View Musical Expression as a Form of Property*, 72 U. CIN. L. REV. 1405, 1411 (2004) (suggesting that "economic considerations supply the justification for copyright law in the United States"); *see also* Jiarui Liu, *Copyright for Blockheads: An Empirical Study of Market Incentive and Intrinsic Motivation*, 38 COLUM. J.L. & ARTS 467, 470 (2014) (noting that "Anglo-American copyright law is widely believed to follow the utilitarian tradition by providing necessary incentives for intellectual creation as a means to promote social

States Constitution reflects this purpose and lays the foundation for federal copyright law "to promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries."[20] The primary basis of American copyright law today was codified in the Copyright Act of 1976 ("Act"), which guaranteed the owners of musical works the exclusive right to reproduce and distribute copies,[21] prepare derivative works,[22] and perform and display copyrighted works.[23] However, the Act also limited the scope of these exclusive rights in order to encourage other artists to develop their own musical works without the fear of lawsuit.[24] In order for a copyright action to succeed, the injured party must prove (1) that the injured party owned a valid copyright of the infringed work, and (2) that the infringing party copied original and protected elements of the work in question.[25]

### *B. Methods of Music Plagiarism Analysis*

American musical tradition exemplifies the "melting pot" phenomenon entrenched in the national collective consciousness, and the widespread appropriation of musical themes from around the world resulted in entirely new genres of music, as well as a society highly concerned with the protection of individual property rights and the importance of borrowing from existing musical libraries to derive novel independent works.[26] It is not surprising, then, that courts have employed a variety of different tests in music plagiarism cases in order to balance these seemingly conflicting objectives.[27]

The "similarity and access" test was a standard employed by the Second Circuit Court of Appeals, most notably in *Arnstein v. Porter*.[28] This test relies on the two elements of copyright infringement—that the "defendant copied from plaintiff's copyrighted work" and "went to far as to constitute improper appropriation."[29] In order

---

welfare" and that "[c]opyright law is widely perceived as the means to promote social welfare by providing a necessary incentive for intellectual creation"). Harper & Row, Publishers, Inc. v. Nation Enterprises, 471 U.S. 539, 546 (1985) (stating that the reason copyrights grant an author a monopoly over a particular work is "in order to benefit the public" and that a copyright's intended purpose is "inducing the creation of new material of potential historical value").

[20] U.S. CONST. art. I, § 8, cl. 8.

[21] 17 U.S.C. § 106(1) (1947); 17 U.S.C. § 106(3) (1947).

[22] 17 U.S.C. § 106(2) (1947).

[23] 17 U.S.C. § 106(4) (1947); 17 U.S.C. § 106(5) (1947); 17 U.S.C. § 106(6) (1947).

[24] 17 U.S.C. § 114 (1947) (exclusive rights to a sound recording "do not extend to the making or duplication of another sound recording that consists entirely of an independent fixation of other sounds, even though such sounds imitate or simulate those in the copyrighted sound recording").

[25] Feist Publications Inc. v. Rural Telephone Service Co., 499 U.S. 340, 361 (1991).

[26] *See Arewa, supra* note 18, at 615-618 (ragtime "set the stage" for jazz and blues music, which exhibit "significant influences" from "African, European, and Caribbean traditions;" other examples include rock and roll borrowing from folk, rockabilly, and African American traditional music and gospel; rockabilly itself was influenced by the Afro-Cuban Habanera genre; the article provides further examples of musical influences throughout).

[27] *See* Iyar Stav, *Musical Plagiarism: A True Challenge for the Copyright Law,* 25 DEPAUL J. ART TECH. & INTELL. PROP. L. 1, 15 (2014).

[28] 154 F.2d 464, 468 (2d Cir. 1946).

[29] *See id.*

to prove the first element of the copyright infringement claim, a plaintiff must show that the defendant had access to the copyrighted work, and that similarities exist between the two works that are "sufficient to prove copying."[30] The plaintiff may show proof of access through circumstantial evidence, and it is often enough to merely show that defendant had an opportunity to view the copyrighted work to create an inference of access for a trier of fact.[31]

If access has been circumstantially proven, even a small amount of similarity might result in a jury trial, but if there is little to no evidence of access, the plaintiff must prove a "striking similarity" between the two works.[32] Striking similarity is a much more complex issue than access and a finding of fact primarily depends on the testimony of expert witnesses.[33] An expert witness will generally single out similar phrases or segments of the works and give an expert opinion claiming that the similarity of a work cannot reasonably be construed as a coincidental independent construction.[34] Although the expert witness presents an objective standard alongside opinion, the striking similarity test still presents a problem of subjectivity, and the court must take into account the probative value of similarities to decide if they are significant or merely common musical themes.[35]

After a showing of both access and similarity, a lay listener (i.e. the finder of fact) must determine whether the appropriation was improper.[36] The improper appropriation test is a primarily economic test; plaintiff simply shows that by copying the part of the music that a lay listener finds appealing, the appropriation of plaintiff's work resulted in a violation of plaintiff's property rights and thus diminished the

---

[30] *See id.*

[31] MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.02[A] (2014) [hereinafter NIMMER]. Widespread popularity of a song is usually enough to create an inference of access. *See* Bright Tunes Music Corp. v. Harrisongs Music, Ltd., 420 F. Supp. 177, 179 (S.D.N.Y. 1976) (concluding that Harrison had access to the copyrighted song for the sole reason that it was number one on the charts in the United States and number twelve in the United Kingdom); Repp v. Webber, 132 F.3d 882, 889 (2d Cir. 1997) (noting that plaintiff need not prove access—if the similarity between the works is strong enough then the trier of fact does not need to make a separate finding for access, as access will be inferred); *cf. Arnstein*, 154 F.2d at 468 (stating that no amount of access may result in a finding of infringement if there is no similarity between the works).

[32] *Arnstein*, 154 F.2d at 468 (noting that without proof of access, "the similarities must be so striking" that there is no possible way the two composers "independently arrived at the same result"); *but see* Selle v. Gibb, 741 F.2d 896, 902 (7th Cir. 1984) (even if the plaintiff proves a striking similarity between the two works, the plaintiff must still also provide a minimal amount of evidence to reasonably infer that the defendant had access to the copyrighted work).

[33] *See Arnstein*, 154 F.2d at 468 (explaining that expert testimony may be offered to assist a trier of fact); *see also* NIMMER, *supra* note 31, at § 13.02[B] (stating that "expert testimony may be necessary to establish striking similarity in 'technical' areas, such as music").

[34] NIMMER, *supra* note 31, at § 13.03[E][3][a][i].

[35] *See* Velez v. Sony Discos, 2007 U.S. Dist. LEXIS 5495, at *30-31 (S.D.N.Y. Jan. 16, 2007) (deciding that not all similarities are probative to the point of proving that the parties could not have independently produced the similarity, e.g., a 4/4 rhythm or a common theme); *see also* McRae v. Smith, 968 F. Supp. 559, 566 (D. Colo. 1997) (stating that a court must be "mindful of the limited number of notes and chords available to composers" and that "common elements" arise in music composition that must not affect a striking similarity determination; the *McRae* court found that because the similar chord progression was one of "the most common chord progressions in all of the music of Western civilization" it could not be used to find striking similarity).

[36] *Arnstein*, 154 F.2d at 468; *See also* NIMMER, *supra* note 31, at § 13.01[B] (appropriation may be nonjusticiable as a matter of law).

"potential financial returns" owed to plaintiff.[37] Because the improper appropriation test depends on the personal feelings of a lay listener and rejects the use of expert testimony, it shares some similarity with the "intrinsic" test described in the next paragraph.[38]

The similarity and access test makes no distinction between objectivity and subjectivity of musical works, but appears to focus on an objective standard for plagiarism litigation by relying on expert testimony.[39] The "extrinsic and intrinsic" test adopted by the Ninth Circuit as an alternative to the similarity and access test divides the plagiarism test into a finding based on a combination of objective and subjective factors, which must prove a "substantial similarity" between the two works.[40] The extrinsic test is intended to be an objective view of the plagiarism case and independent of the conclusions of a trier of fact; therefore the test depends heavily on expert testimony and "specific criteria which can be listed and analyzed," much like the similarity and access test of the Second Circuit.[41] The Ninth Circuit resolved that an idea may be determined by fact, and judged objectively, but an expression of that idea must be judged using different criteria.[42] The intrinsic test examines the expression of an idea and the interpretation of that expression by an ordinary reasonable person.[43] The intrinsic estimation of similarity performed by a lay listener is "virtually devoid" of objective examination and instead focuses on the instinctive reaction to the two different works.[44] Subsequent case law has resulted in the extrinsic test analyzing all objective similarities, whether in idea or expression, and the intrinsic test accounting for the remaining subjective interpretation.[45]

A third test, "abstraction and filtration," is a two-step process in which the court first creates "several levels of abstractions," to examine similarities between the two works to prove "probative similarity," which, like in the Second Circuit, is balanced

---

[37] *Arnstein*, 154 F.2d at 468 (finding that a plaintiff's property interest is not in reputation, but in economic consideration; by copying what makes plaintiff's work "pleasing to the ears of lay listeners," and entices them to patronage, a plagiarizer inflicts not insignificant financial harm on a plaintiff).

[38] *See id.* at 473 (declaring that the improper appropriation test depends on the relationship of the lay listener with the music and that expert musical opinions "are utterly immaterial on the issue of misappropriation"). According to *Arnstein*, if an expert witness were to find striking similarity between two works, but a lay listener was unaffected by the similarities, then the appropriation was not improper or injuring because the infringement did not result in a taking of something of value from the plaintiff. *Id.*

[39] *See* Stav, *supra* note 27, at 15.

[40] The Ninth Circuit first employed the extrinsic/intrinsic test in *Sid & Marty Krofft TV Prods. v. McDonald's Corp.*, 562 F.2d 1157, 1164 (9th Cir. 1977). The *Krofft* court cited the *Arnstein* similarity and access analysis as the basis for the extrinsic/intrinsic test it formulated. *Id.* at 1164-1165.

[41] *Id.* (the extrinsic test must depend on "specific criteria that can be listed and analyzed").

[42] *Id.* (the court compared the extrinsic test to the idea of a nude statue—a nude horse or a nude painting would empirically differ in idea, but not necessarily expression, and may be objectively compared); *but see* Shaw v. Lindheim, 919 F.2d 1353, 1357 (9th Cir. 1990) (explaining that the extrinsic test may be used to analyze both ideas and expressions because the criteria for an extrinsic test "encompass[es] all objective manifestations of creativity"; the *Shaw* court proposed that the extrinsic/intrinsic test should be better understood as an objective/subjective test).

[43] *Krofft,* 562 F.2d at 1164.

[44] *Shaw*, 919 F.2d at 1357.

[45] *See id.*

against evidence of access.[46] The purpose of this abstraction process is to isolate every "constituent structural part" of a work that can be compared, and the levels of abstractions can be imagined as cross-sections of the work, taken from the most objective and quantifiable description, extracted piece by piece down to the most abstract and subjective element of the work that may be reasonably compared.[47] After breaking down the works into abstractions, the filtration process involves taking the collected abstractions and removing "ideas, concepts, and processes," which are unprotected as a matter of law.[48]

### *C. Exceptions to Musical Plagiarism*

All three of the tests above involve some mechanism for the removal of elements that are not eligible for a copyright infringement claim.[49] The "fair use" exception is laid out in the Copyright Act of 1976 and describes explicit purposes in which copying particular elements of a work is not an infringement of that work's copyright.[50] These "purposes include reproduction of a protected work for purposes such as criticism, comment, news reporting, teaching, scholarship, or research," and a court will analyze factors of the copying that contribute to its classification as fair use.[51] Using this model, a parody of a song, for example, would not be considered infringement.[52] Scènes à fair are ideas in a work that cannot be infringed upon because of their necessity or

---

[46] Stav, *supra* note 27, at 18; *see also McRae*, 968 F. Supp. at 562-563 (weighing access and similarity to determine infringement).

[47] *See id.* Judge Learned Hand first imagined the abstraction test in a case in which a playwright sued a motion picture company for adapting his work without credit. Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930). Although there were few similarities in the physical content of the works, including marked differences between the text, characters, and plot, the two works appeared to share a common theme of "a quarrel between a Jewish and an Irish father, the marriage of their children, the birth of grandchildren and a reconciliation." *Id.* at 122. If plagiarism was not addressed accounting for abstract concepts "a plagiarist would escape by immaterial variations." *Id.*; *see also* Computer Assocs. Int'l v. Altai, 982 F.2d 693, 706 (2d Cir. 1992) (stating than an abstraction-filtration-comparison test involves breaking down a work into its "constituent structural parts").

[48] *Computer Assocs. Int'l*, 982 F.2d, at 707.

[49] *See supra* Part II.B discussing improper appropriation in the similarity and access test, and how it limits the infringement to the especially unique elements of a claimant's work, which are pleasing to the ears of a lay listener. *See supra* Part II.B discussing the extrinsic test and how it involves the removal of nonjusticiable elements. *See supra* Part II.B. discussing the filtration step of the abstraction and filtration test and how it separates abstractions that are "unprotected as a matter of law."

[50] 17 U.S.C. § 107 (1976).

[51] *See id.* (detailing the factors which determine fair use, which include "(1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work").

[52] *See* Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 571 (1994) (holding rap group 2 Live Crew's parody of Roy Orbison's "Pretty Woman" to be considered fair use).

commonplaceness within a particular genre or medium.[53] A "de minimis" exception occurs when the copied part is too insignificant to amount to plagiarism.[54]

## III. ANALYSIS

### *A. Limitations of Expert Testimony*

The critical complication involved in all three of the tests described above is the difficulty in separating the objective from the subjective.[55] Particularly, the extrinsic/intrinsic test illustrated in Part II suffers from inappropriate application of law due to mingling of objective and subjective analysis in the extrinsic phase of the test.[56] Courts throughout the United States liberally apply evidence proffered by expert testimony during an analysis of similarity between two works.[57]

However, because the tests for copyright infringement often involve a dual natured test involving objective evidence presented by an expert and a subjective test as interpreted by a lay listener, courts must carefully determine what evidence provided by an expert should be admitted as "objective."[58] A 2004 Ninth Circuit case,

---

[53] *See* Alexander v. Haley, 460 F. Supp. 40, 45 (S.D.N.Y. 1978) (ruling that there are "incidents, characters or settings which are as a practical matter indispensable, or at least standard, in the treatment of a given topic"); see *also* Stav, *supra* note 27, at 20 ("For example, an instrumental introduction in a song or the main character's last words before his death in a movie easily qualify for a scènes à faire defense."); *Swirsky,* 376 F.3d at 850 (Ninth Circuit district court ruled that the Mariah Carey song "Thank God I Found You" was not violative of plaintiffs' "One of Those Love Songs," on the basis that the melody was reminiscent of the classic and unprotected song "He's a Jolly Good Fellow;" however, the court of appeals reversed this ruling—this reversal highlights the difficulties of ruling on scènes à faire and extrinsic analysis). Scènes à faire analysis requires that the court examine whether "motive" similarities that plaintiffs attribute to copying could actually be explained by the common-place presence of the same or similar "motives" within the relevant field. *Id.*

[54] "De minimis" comes from a Latin expression literally meaning "the law does not concern itself with trifles." *See* Stav, *supra* note 27, at 20; *see also* James W. Newton v. Diamond, 204 F. Supp. 2d 1244, 1256 (C.D. Cal. 2002) (stating that an "alleged infringer must demonstrate that the copying of the protected material is so trivial" that the copying does rise to the "quantitative threshold of substantial similarity"). "[S]mall and unoriginal portions of music cannot be protected by copyright." *Id.* The de minimis exception appears to bear some similarity with the improper appropriation test of the Second Circuit and the intrinsic test of the Ninth Circuit, which rely on the average lay listener's natural response to the music; *cf Newton v. Diamond*, 349 F.3d 591, 594 (9th Cir. 2003) (declaring that "use is de minimis only if the average audience would not recognize the appropriation").

[55] *See Swirsky,* 376 F.3d at 848 (relying on subjective opinions of the expert witness to influence the objective aspect of the plagiarism determination); *see also supra* Part II.B ("Although the expert witness presents an objective standard alongside opinion, the striking similarity test still presents a problem of subjectivity.").

[56] The extrinsic analysis is intended to be entirely objective, but because true objectivity is impossible there is always a risk of subjectivity corrupting an extrinsic test.

[57] *See* Three Boys Music Corp. v. Bolton, 212 F.3d 477, 485 (9th Cir. 2000) ("The extrinsic test often requires analytical dissection of a work and expert testimony."); Positive Black Talk, Inc. v. Cash Money Records, Inc., 394 F.3d 357 (5th Cir. 2004) (Fifth Circuit relied on expert testimony for a substantial similarity finding); Cottrill v. Spears, 2003 U.S. Dist. LEXIS 8823, at *32 (E.D. Pa. May 22, 2003) (Third Circuit district court determined substantial similarity through expert testimony).

[58] *See Arnstein*, 154 F.2d at 468 (explaining that expert testimony should be admissible for analysis of musical similarities but not as to "illicit copying," which should be left to a lay listener).

*Swirsky v. Carey*,[59] highlights the problems that a court may have when determining extrinsic evidence and the limitations of a standard dependent on separating "objective" and "subjective" into "assisted by expert testimony" and "unassisted by expert testimony" respectively.[60] Ninth Circuit courts have long claimed that the extrinsic test concerns only objective aspects of a work, but the reliance on experts to teach lay listeners how to objectively appreciate the music muddles the distinction between objective and subjective.[61] The exceptions to a music plagiarism finding, primarily scènes à faire and de minimis, depend heavily on an expert's testimony that a musical element is sufficiently common or minor enough to not constitute plagiarism—because an expert's testimony is admitted in order to assist a trier of fact to better understand an issue of which the trier does not have special knowledge.[62] Courts have rejected the idea of a similarity test based entirely upon a purely scientific analysis of the constituent elements of a work.[63] However, courts have adopted no coherent model to determine substantial similarity, and appear to have stuck with the idea that music is too ethereal, or simply consists of too many elements to be effectively categorized or compartmentalized and separated into individual elements.[64] In the *Swirsky* case, for instance, the fact finder relied primarily on the expert testimony of a musicologist, who selectively omitted notes which he subjectively described as "ornamental" to a listener.[65] The court allowed the expert's subjective testimony under the assumption that a musicologist has a better understanding of musical similarity and the underlying principles of music than a lay-listener, who is less equipped to comprehend objective similarity, so the removal of ornamental notes to show comparison was considered in the context of helping to assist in the objective determination of the extrinsic test.[66] Despite the belief that experts are more capable of listening to and understanding music than an average listener, "there is no scientific evidence that 'expert listeners' are any better than lay listeners for forensic challenges

---

[59] 376 F.3d 841 (9th Cir. 2004).

[60] *See id.*

[61] *See Krofft*, 562 F.2d at 1164.

[62] Fed. R. Evid. 702.

[63] *Swirsky,* 376 F.3d at 847-48 ("Objective analysis of music under the extrinsic test cannot mean that a court may simply compare the numerical representations of pitch sequences and the visual representations of notes to determine that two choruses are not substantially similar, without regard to other elements of the compositions."); *see also* Paul W. Orth, *The Use of Expert Witnesses in Musical Infringement Cases*, 16 U. PITT. L. REV. 232, 235 (1955) (This article from the 1950s, written long before massive advancements in the abilities of computer programs, relates that a purely analytical discussion of music comparison must lead to false positives for musical plagiarism or otherwise not totally capture the whole of the issue).

[64] *See, e.g., Swirsky,* 376 F.3d at 849 (acknowledging that there is no finite set of categories and factors to analyze music plagiarism and that the court would not proceed to establish any sort of specific guidelines in the current case).

[65] *See id.* at 846.

[66] *See id.* at 847 (explaining that a court may analyze only the parts of a song relevant to appropriation and remove "all the elements unique to [a plaintiff's] performance;" what parts should be removed and whether they constitute plagiarism remain largely in the hands of expert witnesses); *See generally* Begault, *supra* note 16, at 4 (the belief that musicologists are better equipped for listening to music versus the lay listener due to training and expertise may be known as the "golden ear" rule); *see also* M. Fletcher Reynolds, *Selle v. Gibb and the Forensic Analysis of Plagiarism*, 32 C. MUSIC SYMP. 55, 59 (1992) (because judges have little understanding of musicology or a musicologist's methodology, courts often do not "steer" a musicologist's "evidentiary decisions in the right direction" and due to lack of expertise may allow "truly astonishing misinformation").

such as speech intelligibility from noisy recordings, or presumably, for determining whether or not there is a substantial similarity between two recordings."[67]

The "inherent complexities of music" and the difficulty of categorization present a unique challenge to music analysis in a copyright case.[68] A commonly used objective metric is "note-for-note comparison," where experts draw lines between notes in the two works that have the same pitch, sometimes transposing the music to display the most similarity between the melodies of the two compositions.[69] This approach may be misleading to a jury and is decidedly unscientific, particularly when access to more robust similarity measures are available.[70]

The solution, of course, is not to discount all expert testimony, but perhaps computational models can assist music plagiarism litigation to a greater extent than in the past.[71]

### *B. The Construction of the Musical Plagiarism Model*

Software for music comparison has existed for a long time, but has rarely seen practical application in the sphere of law regarding plagiarism disputes.[72] This is, at least in part, due to an absence of musical comparison software specifically tailored to legal applications, as the majority of audio comparison software exists to appeal to commercial consumers or other, non-legal, areas of academia.[73] However, some universities, as well as private companies, have recently endeavored to extend the purview of music comparison software to deal with the complex legal issues of copyright.[74]

---

[67] Begault, *supra* note 16, at 4 (citing F. Poza & D.R. Begault, *The role of transcriptions in the courtroom: A scientific investigation*, 33rd Audio Engineering Society Convention (2008)).

[68] Robert J. S. Cason, *Singing from the Same Sheet: Computational Melodic Similarity Measurement and Copyright* Law, 26 INT'L REV. L. COMPUTERS & TECH. 25, 27 (2012) (assessing similarity in music is difficult because of the "inherent complexities of music," because context is important to a plagiarism determination, and because music is a "nonverbal domain"). *See supra* Part II for discussion of the tests examined that either try to break down music into its constituent elements, as in the abstraction-filtration model, or split deliberation between elements assisted by experts and the non-quantifiable "feeling" of music (extrinsic/intrinsic test) or the similarity and access/improper appropriation test.

[69] *Id.* at 27-28.

[70] *Id*. at 28.

[71] An analysis of current computational models appears *infra* Part III.B.; a proposal for the models' inclusion in copyright law appears *infra* Part III.

[72] Cason, *supra* note 68, at 28. The author of this comment formerly worked as a sight-singer for the audio comparison "query by humming" program *Tunebot*, which can be accessed at http://tunebot.cs.northwestern.edu/ (last visited Dec. 21, 2015).

[73] Cason, *supra* note 68, at 28; *see generally* Michael A. Casey, "SOUND• CLASSIFICATION AND SIMILARITY." INTRODUCTION TO MPEG-7: MULTIMEDIA CONTENT DESCRIPTION INTERFACE 322—223 (2002) (a book chapter dedicated to probabilistic models for audio classification and similarity relates that potential applications of the model described within "will be valuable components in new Internet music software, professional sound-design software, composers tools, audio-video search engines and many yet-to-be-discovered applications;" the article makes no mention of potential copyright comparison applications).

[74] Cason, *supra* note 68, at 28; Fraunhofer Institute for Digital Media Technology offers a music plagiarism detection software "to detect both sampling plagiarism and melody plagiarism in music recordings." *See Music Plagiarism Detection*, FRAUNHOFER IDMT,

Music plagiarism software may be utilized primarily to assist in two areas: (1) to contribute to the objective determination of musical similarity between two works in a copyright case; and (2) to help prevent a songwriter from potentially infringing on a similar work during the production process before the release of a song.[75]

Of course, the genre and context of a work play a major factor in a determination of similarity, which suggests that expert testimony is necessary to elucidate the factors that may be unique to a certain style of music or whether a musical theme can be described as a scènes à faire.[76] Genre, however, is a very difficult concept to define, and many bands or songs may fall under multiple genres or sub-genres.[77] Technology may be able to assist experts in a determination of genre by employing machine learning algorithms, which may use a database of songs classified by genre, analyze the components of each individual work, and prescriptively classify further audio into one or more of the genres supplied by the database.[78] Therefore, when Robin Thicke says he wants to make a song in the "groove" or style of Marvin Gaye, but does not want to fall victim to a copyright infringement lawsuit, a machine learning algorithm might be able to model how closely the song fits into the Motown genre or if it more resembles Marvin Gaye's style than another artist's.[79] This machine learning methodology is still subject to problems of classification, as it requires manual input to provide the categories that the algorithm processes. Thus, in order to see if a song more closely resembles either Rock or Hip-Hop, a researcher must provide a database of songs already classified under "Rock" or "Hip-Hop." In other words, the process which determines the genre of the songs already classified is up to the researcher.[80]

The reverse of the machine learning methodology described above involves creating a genre taxonomy from the ground up out of a database of music by grouping

---

http://www.idmt.fraunhofer.de/en/institute/projects_products/m_p/music_plagiarism_detection.html (last visited Dec. 21, 2015). The music plagiarism analyzer tool developed by IDMT in cooperation with the REWIND project may be downloaded at https://sites.google.com/site/rewindpolimi/downloads/tools/audio-music-plagiarism-analyzer (last visited Dec. 21, 2015).

[75] *See id*. (Fraunhofer IDMT suggests that its software can be used to help "independent reviewers, musicologists, composers, music labels, or publishers in detecting cases of music plagiarism using objective criteria.")

[76] *See Swirsky*, 376 F.3d at 850 (where the difference in genre between two works played a role in whether a phrase could be considered a scènes à faire).

[77] Arewa, *supra* note 18 and accompanying text. *See supra* Part II.B. Many bands borrow from multiple styles of music in their work so genre classification is a particularly difficult task, even for experts. *See generally* Jean-Julien Aucouturier & Francois Pachet, *Representing Musical Genre: A State of the Art*, 32 J. NEW MUSIC RES. 83 (2003) ("Genre is intrinsically ill-defined and attempts at defining genre precisely have a strong tendency to end up in circular, ungrounded projections of fantasies.").

[78] *See id.* at 84-90. The Aucouturier article observes eight different machine learning algorithms. *Id.* All of the algorithms begin with a preexisting database of songs coded into a genre. *Id.* The genres imposed upon the algorithms are of the researchers' determination, and may not necessarily be musical genres, e.g., jazz, funk, rock, but could be classified by artist or any number of schemata. *Id.* The audio is then broken down, frame by frame, and a value applied to certain elements that may make up a genre. *Id.* In this case, the algorithms examined timbre, rhythm, and pitch. *Id.*

[79] *Williams,* 2015 U.S. Dist. LEXIS 97262, at *36; *see also* Aucouturier & Pachet, *supra* note 77 (explaining that classification may be done by artist, genre, or any number of arbitrary variables endogenous to the metrics of the algorithms).

[80] *See id*. (explaining that the taxonomies implemented in the machine learning models are arbitrary).

titles according to their similarity.[81] The computation of this reverse model may utilize the same compositional variables which govern the machine learning algorithms' analyses, such as timbre, rhythm, and pitch, though not necessarily.[82] Where the previous methodology sorted songs into genre classifications already provided, the reverse model instead applies similarity metrics in order to cluster the audio tracks into classifications based on degree of similarity.[83]

A number of musical databases with different methods of advanced classification already exist, which may mitigate the problems of accuracy of an input database fed into a machine learning algorithm, or provide valuable data for similarity comparison based on the variables comprising the database classifications.[84] For example, internet radio company Pandora Media Inc. sorts the songs in its database using hundreds of different musical characteristics, so an expert witness in a copyright case may be able to present stronger objective evidence of similarity using a database such as Pandora's in conjunction with the expert's own knowledge.[85] Because Pandora's database is managed by a team of experts rather than a single musicologist's opinion, statistical biases of subjectivity can be minimized.[86] A database created specifically for music copyright law could allow for rigorous analysis of each work individually by categorizing a song according to all of its compositional elements.[87] Because each individual composition is first analyzed according to an objective scale and then compared, there is no potential for bias from a musicologist who begins an analysis by directly comparing two songs that are currently in controversy.[88] Furthermore, the two works can be compared to other songs within the same genre to provide an idea of

---

[81] *See id.* at 90.

[82] *See id.* Distance functions applied in the machine learning model may be used to "assess similarity between individual titles," but the models which attempted this type of assessment "are on the edge of genre classification," and did not explicitly attempt to build genres out of a similarity metric. *Id.* Furthermore, "intrinsic signal attributes" such as timbre, may not necessarily be correlated with traditional genres. *Id.*

[83] *See id.* at 90–91. A "Collaborative Filtering" model may help group songs into similar genres without relying on a song's compositional elements. *Id.* This model relies on the preferences of users and grouping those preferences with those of similar users. *Id.* Another method of categorization looks for genre similarity between two songs by how often they appear near each other, such as on radio stations or in music playlists or compilation albums. *Id.*

[84] *See* Yvette Joy Liebesman, *Using Innovative Technologies to Analyze for Similarity Between Musical Works in Copyright Infringement Disputes*, 35 AIPLA Q. J. 331, 356–357 (2007) (explaining the process in which Pandora Media Inc. builds its music database of over 10,000 artists, producing "qualitative, objective analyses" of songs; a music expert listens for "up to 400 distinct musical characteristics" and distributes them "into larger categories such as harmony, rhythm, structure, melody, vocals, and lyrics;" the article also mentions another, less rigorous, database known as "The Global Jukebox").

[85] *See id.*

[86] *See id.* at 349 (multiple experts can also minimize subjective bias across a database by classifying the same song multiple times with different experts).

[87] *See id.* (an analytic test involving categorization of a large number of variables of a work "is a reasonable and viable expansion of the current music copyright infringement test"); *but see* Cason, *supra* note 68, at 31 (doubting that such a database, which requires extensive manpower and monetary investment, is practical without commercial incentive, and doubting that a similarity analysis provided by such a database would "include melodic feature-descriptors that are fine-grained enough to allow for valid similarity measurements at the level of detail required to aid cases of plagiarism").

[88] *See id.* at 348-349.

the average level of similarity common to the genre.[89] This sort of analysis may be beneficial for proving objective similarity in a similarity and access test or in the extrinsic segment of a Ninth Circuit test, and because the analysis involves segmenting each relevant aspect of a composition it may be of particular use in an abstraction-filtration test.[90]

In the past, the processing ability of technology and data storage capabilities severely limited the prospect of analysis using large music databases driven by an extensive list of variables.[91] Since 1980, the world's per-capita ability to store data has about doubled every 40 months, so the rapid speed of technological advancement has surely outpaced the legal sphere, which has yet to embrace many methods of computational analysis.[92]

Robert Cason and Daniel Müllensiefen recently introduced a compelling new model for a computational similarity measurement that may help in a copyright infringement dispute.[93] Like the previous models, this statistical similarity method needs a database, but this model only requires musical notation because it solely measures the similarity of monophonic melodic elements of a work.[94] This model, instead of focusing on the similarities between two works as a whole, concentrates on melodic similarities between two works that are *not* held in common with other songs.[95] While emphasizing the similarities that are unique to the two works, the model can account for scènes à faire by placing a lower statistical weight on melodic elements that are common to all of the songs in the database.[96]

In practice, the model successfully predicted, with 90 percent accuracy, the outcome of a set of 20 U.S. copyright cases where the subject of the infringement was melodic similarity.[97] The Cason and Müllensiefen system may be easily implemented into a copyright infringement case, but its relegation to only monophonic melodic similarity limits its practicality to only a subset of copyright infringement cases.[98]

---

[89] Cason, *supra* note 68, at 31.

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

[91] *See* BEN M. CHEN, HARD DISK DRIVE SERVO SYSTEMS 7 (2006) (in the year 2000, the largest hard disk drive on record was 180 GB).

[92] Martin Hilbert, *The World's Technological Capacity to Store, Communicate, and Compute Information*, 332 SCIENCE 60, 64 (2011); *see also* Depoorter, *supra* note 18 (copyright law often lags behind technological advancement due to legal uncertainty, resulting in legal delay).

[93] Cason, *supra* note 68, at 32.

[94] *Id.* Because the Cason and Müllensiefen model does not require a large amount of human input for classification, such as the Liebesman model *supra* note 84, it is a considerably cheaper and practicable method. *Id.* at 31.

[95] *Id.* The model draws from a database of 14,063 transcriptions of MIDI pop songs. *Id.* at 33. Each song is broken down into short melodic phrases, with each phrase weighted according to its commonality, where more unique phrases have a higher weight.

[96] *Id.*

[97] *Id.* at 34.

[98] *See Swirsky,* 376 F.3d at 848 (commenting that a court cannot assess similarity based solely on one element, without taking the whole picture into account); Cason, *supra* note 68, at 35. The Cason and Müllensiefen model mimics the work of an expert witness by highlighting unique similarities in the works and eliminating similarities common to the database. *Id.* The experiment performed in the Cason article used a database of only pop songs, and there could be significant differences in the results with different genres, which may contain more or less complex melodic elements, or there may be a song of one genre infringing upon a song of a different genre. *Id.* The test can only account for singular notes in succession, and would be ineffective if the similarities involved chords or harmonies. *Id.*

Polyphonic similarity, where multiple notes are played in unison, involves a more complex system of analysis than the Cason and Müllensiefen similarity test.[99] A novel polyphonic plagiarism detection model introduced this year resembles the machine learning methodology described previously in this section by segmenting a work into frames and classifying the similarity between two works using objective audio features such as timbre, rhythm, or pitch.[100]

All of the computational models described above paint an optimistic picture for the future of objective analysis in copyright law, either to assist expert witnesses or to prevent an artist from infringing in the first place.[101] However, these computational models are plagued by potential unique challenges of practicality, efficacy, or particularity; adopting them into the current tests for music plagiarism may be a slow and arduous process.[102] Perhaps if courts were to apply an objective test more favorable to statistical analyses, computational models would be more capable of providing objectivity to copyright litigation.[103]

## IV. PROPOSAL

A proposal by Professor Iyar Stav suggests expanding Cason and Müllensiefen's concept from a weighted statistical melodic comparison model into a "Full Statistical Test."[104] This Full Statistical Test is the basis of my own proposal and closely resembles the arbitration-filtration test, but focuses on a purely quantitative determination.[105] This test requires splitting the numerically comparable elements of two works into their component parts and finding the respective similarities between

---

[99] Soham De, *Plagiarism Detection in Polyphonic Music using Monaural Signal Separation,* arXiv.org preprint (Feb. 27 2015), available at http://arxiv.org/pdf/1503.00022v1.pdf.

[100] *Id.* at 4. "[T]raditional representations for audio," such as timbre, rhythm, or tonality are extracted from the audio, followed by non-negative matrix factorization based features. *Id.* The two works are then compared in similarity according to those extracted features. *Id.*

[101] *See* Cason, *supra* note 68, at 35 (claiming that their computational system is intended to "model the reasoning and decisions" of expert witnesses in order to "guide and inform human judgments"); *see also* Liebesman, *supra* note 84, at 349 (Liebesman's comprehensive public database, if constructed, could assist experts in a copyright infringement case and prevent end users from accidentally infringing on a copyright); Fraunhofer IDMT, *supra* note 63 (Fraunhofer IDMT's plagiarism software is marketed to "independent reviewers, musicologists, composers, music labels, [and] publishers").

[102] *See* Cason, supra note 68, at 31 (proposing that it is unlikely for Liebesman's database to get funded, and that if a private entity were to fund such a database, it would have no incentive to publicize the database for analytical copyright usage). The Cason and Müllensiefen model can only analyze the similarity between melodic elements, and only then using monophonic notation. *Id.* at 35; *see also* De, *supra* note 99, at 5 (plagiarism usually involves only a small amount of copying, and the polyphonic model has no specific mechanism to address this, unlike the Cason and Müllensiefen model).

[103] Stav, *supra* note 27, at 50-52. A proposal for a more scientific objective test will appear *infra* Part IV.

[104] *Id.* at 50-51; *see also* Cason, *supra* note 68 (the Cason and Müllensiefen model values the statistical weight of a melodic similarity by judging whether a melodic phrase is common to other songs of the same genre).

[105] Stav, *supra* note 27, at 50-51 (explaining that the test divides the comparable musical elements, such as melody, rhythm, or tempo, into separate cells, analyzing them independently. *See, e.g.*, Part II discussing the abstraction process.

the quantitative elements of the two works.[106] Then, each element is weighted in statistical significance according to its commonality among all musical works in a database.[107] Comparing these two statistics, the model creates a value for plagiarism likeliness dependent on the similarity value and the commonality value.[108]

Stav suggests that this method could be used to add legitimacy to the proposals of an expert witness.[109] But this method can be augmented for greater accuracy and given more responsibility in a court of law than Stav suggests.

First, the test described above can be improved by treating a combination of variables as a single element to plug into the model, rather than adding the values separately to suggest plagiarism.[110] This can make a stronger argument, particularly if the expert thinks that the plagiarism is due to a combination of the musical elements, and the individual elements themselves do not surpass commonality.[111] The test also needs to take into account the fact that certain elements do not constitute a reliable estimation of plagiarism even if their output value is high.[112] This type of false positive may still need to be evaluated by musical expert opinion, or the model could place a lower weight on elements that are less likely to provide a meaningful estimation of plagiarism even with a higher value according to the model, such as time signature.[113] The inherent concern with adding more variables to the model is that each additional variable increases the chance of finding some sort of positive value, even if the similarity/commonality value is negligible to a finding of plagiarism due to the nature of the variable.[114]

---

[106] *See id.* at 51 (stating that "the similarities between each couple of parallel cells will . . . be summarized with a numerical value").

[107] *See id.* ("Each cell in this grid is analyzed separately in comparison with the corresponding cells of all the other songs." "With this analysis . . . the structure of each cell is summarized with a statistical value that represents the probability of composing an identical part").

[108] *See id.* ("The general statistical value of the complaining song is then placed in an equation with the two songs' similarity count; a combination of low general statistical value and high similarity count indicates a high probability of copying and vice versa.").

[109] *See id.*

[110] For example, perhaps melody produces a low output value due to commonality and tempo produces a low output value due to commonality, but a variable combining the melody and tempo together generates a more significant output value. Mathematically, both similarity and commonality are necessarily lower in the combination variable; *cf. Swirsky*, 376 F.3d at 848 ("[T]o disregard chord progression, key, tempo, rhythm, and genre is to ignore the fact that a substantial similarity can be found in a combination of elements, even if those elements are individually unprotected").

[111] A problem with this test that cannot be easily eliminated is that certain elements of a musical work stand out more to listeners as plagiarism, and it is ultimately in the hands of the jury to decide whether there is a certain level of similarity between the songs. *See* Stav *supra* note 27, at 52 ("However, the final decision still must originate through the untrained ears of lay listeners."); *cf.* Aucouturier & Pachet *supra* note 79, at 89-90 (recognizing that two songs of completely different genres may have very similar timbres, e.g. "a Schumann sonata ('Classical') and a Bill Evans piece ('Jazz')." This might lead to a false positive in a statistic test when a listener would easily discern there was no plagiarism.

[112] For example, a 17/4 time signature is very rare but in no way should constitute plagiarism; likewise, a digeridoo is rarely used in popular music, but two bands using a didgeridoo cannot justify a plagiarism finding from the unique timbre alone. This lends credence to the belief that a music plagiarism analysis must include as many factors as is feasible.

[113] *See supra* note 112 and accompanying text.

[114] *Id.*

Also, the test can be fortified by including qualitative metrics in the style of Pandora Media Inc., which was described in the previous section.[115] Surely there will be overlap between the quantitative and qualitative metrics, since the qualitative analyses also include categories for quantifiable factors "such as harmony, rhythm, structure, [and] melody."[116] If this test were to include a database of measurements such as the 400+ variables that Pandora categorizes, it would be easy to do an analysis on commonality and similarity simply by looking at the number of qualitative variables the two works have in common and judging them against the probability of those variables appearing together in other songs. If two works have more elements in common than expected, there is a higher probability of plagiarism. Additionally, Pandora's current proprietary software runs its own mathematical equations to calculate a correlation between songs, which could be used similarly in conjunction with this model.[117] Clearly, a qualitative analysis would have trouble detecting plagiarism by itself, as the current system is optimized only to find songs and artists with similar styles, but if used in conjunction with the quantitative aspects of this statistical test it could provide a stronger justification for a position in a plagiarism dispute.[118]

Daniel Müllensiefen has previously used melodic similarity modeling to predict, with up to 90% accuracy, the outcome of a plagiarism case dependent primarily on melodic copyright infringement.[119] Melody is often a major factor in an infringement case, but it is rarely the only factor, and adding additional relevant similarity measurements could close the gap in predictive accuracy or better explain if other variables caused the case to turn a certain way.[120]

I also suggest applying Stav's full statistical test to previous music copyright infringement cases in order to create a predictive model to better understand how similarity of certain variables affect the outcome of a music copyright case. A statistical regression model, for example, could produce an estimate for how much the similarity/commonality value of a variable changes the probability of a plagiarism ruling.[121] A problem with this kind of analysis is that there may be a lack of past

---

[115] *See* Liebesman, *supra* note 84, at 346–349 (describing the strategy Pandora's musical experts use to classify songs according to qualitative musical elements within the song, e.g., harmonic structure, gender, vocal range, key tonality, and instrumentation). *See supra* Part II describing the Pandora model and the Liebesman MEA model; *see also supra* note 84, at 347.

[116] Liebesman, *supra* note 84, at 347.

[117] *See id*. (explaining that Pandora's measurements are "stored in a computer database where a proprietary software program mathematically calculates a correlation between songs).

[118] *See id*. ("The database is currently used to allow listeners to find music they like and to suggest songs that are similar to previously suggested music.").

[119] Daniel Müllensiefen, *Court decisions on music plagiarism and the predictive value of similarity algorithms*, 13 MUSICAE SCIENTIAE 257, 275-276 (2009). In the similarity analysis, the highest scoring metric evaluated how much of the relevant part of plaintiff's work was incorporated by the defendant. *Id*. This proved more successful than evaluating over either unions or intersections of the two works. *Id*.

[120] *See id*. at 285 (explaining how the melodic similarity estimate could potentially be combined with other similarity measurements to produce more reliable results).

[121] A logit or probit regression model could explain how much the similarity/commonality measurement of a variable affects the probability of a plagiarism finding. *See* Shayrn O'Halloran, *Lecture 9: Logit/Probit*, Econometrics II, available at http://83.143.248.39/STUDENTS/PVS110/ECO311/probit-logit/Lecture_9_LOGITPROBIT.pdf (last visited Nov. 20, 2015). Logit and probit regressions treat a binary variable (0=no plagiarism, 1=

observations for reliable numerical data, as music copyright infringement cases are uncommon and will also need to be grouped according to the type of plagiarism that has occurred so the variables do not incorrectly suggest an effect or lack of effect.[122] Still, applying just this basic analytical tool to the full statistical test could shine some light on the threshold similarity/commonality value for a copyright infringement case.

A bolder approach for using the full statistical test is to treat it as a barrier to entry for a plagiarism suit, or for courts to grant summary judgement more freely when a plaintiff fails to meet a certain threshold for similarity. Because of copyright law's "low threshold of eligibility," claims with "implausible cases of misappropriation" are brought before the court.[123] Only a few years ago, a copyright infringement claim against Elton John managed to reach the 7th Circuit Court of Appeals.[124] The plaintiff in that case claimed that a combination of lyrical elements amounted to plagiarism, including "a theme of impossible love," "description of the beloved's light eyes," and a three syllable Russian name beginning with the letter N and ending with the letter A.[125] Although the court of appeals ultimately dismissed the suit, the mere fact that the 7th Circuit gave enough respect to this claim to accept it for appeal illustrates the problem of courts treating music as arcane and allowing frivolous music plagiarism cases to survive past a preliminary stage.[126] This leads most copyright infringement defendants to settle out of court, likely to avoid the fate of Pharrell Williams and Robin Thicke in a possible jury verdict.[127]

If courts, however, were to require a certain low-level finding of objective similarity/commonality according to a full statistical test before approaching the next stage of litigation, less judicial resources would have to be wasted to adjudicate claims that truly have no objective similarity proximate to plagiarism. Even if courts were to not apply a concrete requirement, empirical data from music plagiarism models could allow courts the justification to dismiss more freely "implausible cases of misappropriation."[128]

---

plagiarism) as a continuous probability function from zero percent probability to 100 percent probability. *Id*. The probability of a plagiarism finding is dependent on one or more independent variables, which in this case might be the similarity/commonality measurement for melody, rhythm, or harmonies. The regression then can predict how much a change in the similarity measurement of one of the independent variables affects the probability percentage of the dependent variable.

[122] *See Case List*, MUSIC COPYRIGHT INFRINGEMENT RESOURCE, http://mcir.usc.edu/cases/Pages/default.html (explaining that there are less than 200 federal plagiarism opinions since the 1850s; the law has changed considerably over more than a century, different circuits apply different tests, and the type of case differs by the kind of infringement asserted).

[123] Charles Cronin, *I Hear America Suing: Music Copyright Infringement in the Era of Electronic Sound*, USC Legal Studies Research Papers Series No. 14-6, Mar. 14, 2014, at 5, at http://mcir.usc.edu/documents/Croninarticle.pdf.

[124] *See* Hobbs v. John, 722 F.3d 1089 (7th Cir. 2013).

[125] *Id.* at 1094.

[126] *See* Cronin, *supra* note 123, at 4.

[127] *See id*. at 6.

[128] Cronin, *supra* note 123, at 5.

## V. CONCLUSION

The technological software and modeling described previously could help courts grapple with the objective and subjective aspects of musical plagiarism. Although both the judicial system and technological plagiarism models have a long way to go before seeing practical legal results of these analytical tools, the future is promising, as we continue to see a blurring of the lines between objective and subjective analysis.[129]

[129] *See supra* Part III describing the polyphonic plagiarism model. This model was presented just this year, and only as a preprint, but it paves the way for further research into statistical analysis of legal disputes involving musical works. De, *supra* note 99. Many of these models are only prototypes or conjectural. *See* Liebesman, *supra* note 84, at 349 ("The MEA test is a reasonable and viable expansion of the current musical copyright infringement test and should be developed for use by courts"); *see also* Cason, *supra* note 68, at 31 (opining that the any sort of database focused qualitative musical test á la Pandora for a court system is far from any sort of feasible realm of application at this point in time). Cason and Müllensiefen admit themselves that their current model is primarily a demonstrative prototype and limited in application, but provides a building block for further research into musical plagiarism litigation analysis. *Id.* at 35.