Cornell Law Review

Volume 37
Issue 2 *Winter 1952*

Article 5

Re-Examining the Traditional Legal Test of Literary Similarity A Proposal for Content Analysis

Robert C. Sorensen

Theodore C. Sorensen

RE-EXAMINING THE TRADITIONAL LEGAL TEST OF LITERARY SIMILARITY: A PROPOSAL FOR CONTENT ANALYSIS

*Robert C. Sorensen**
*Theodore C. Sorensen**

If the object of copyright laws is to secure to authors for a limited time the *exclusive* use and benefits of their creative works and thus provide them with incentive to further effort, the protection against infringement to-the-point-of-similarity must be carefully enforced.[1]

Seldom can direct proof of copying be introduced. Consequently the author who claims infringement must undertake the proof of four fundamental elements: (1) the existence of his copyright; (2) the fact that his work was copyrightable, being original and not in the public domain; (3) the fact that the alleged infringer had access to the copyrighted work; and (4) finally, evidence of similarity sufficient to indicate an infringement of the copyright without further proof of intent.

The problems of proof involved in the first three elements, except in those cases of disputed originality, offer little challenge to the court in that they depend upon events (often stipulated) the existence of which can be clearly tested by resort to specific law and fact. But proof of similarity requires evidence the validity of which to date is extremely illusive and subject to as many disputed testimonials as the contestants may care to muster. Thus, it is with proof of similarity in an area of constant disputation that copyright attorneys are constantly concerned.

The history of the concept of copyright tells the story of changing functions of the state, technological progress in mechanical fields of communication, and new social ethics. The advantages which may accrue to a writer from today's mass distribution of literary output inspire a somewhat different outlook than was held during the early middle ages when much of the world's literature was disseminated by schools and monastaries of the Church and when "books were composed and classics copied not for profit but for the glory of God and the Church, and the freer the diffusion of religious truth the greater the glory."[2]

It has been no easy matter to detect ". . . those literary cooks, who

* See Contributor's Section, Masthead, p. 718, for biographical data.

1 17 U.S.C. §§ 1, 101 (1950); U. S. CONST., Art. I, § 8, cl. 8; BALL, THE LAW OF COPYRIGHT AND LITERARY PROPERTY 319 (1944). For infringement regardless of intent, also see BALL, at 323, 329.

2 Whipple, *Copyright*, 4 ENCYC. SOC. SCI. 401 (1931).

skim the cream of others' books."[3] Literary plagiarism is a fine art which reveals an enlightened development of more than 20 centuries.[4] It is no mean task to stake off individually owned plots in a universe of discourse—the common denominator known as communication by which human beings seek to understand and utilize each other.[5] Judge Story maintained that every author since the Greek scholars necessarily borrows from his predecessors, since "there are, and can be, few if any, things which in an abstract sense, are strictly new and original throughout."[6]

Whether the evidence is for the persuasion of a judge or a jury,[7] and whether the infringement action is at common law or by statute,[8] it is important for the effective application of copyright law that a test of similarity be designed which protects all parties. Because of the above mentioned difficulties, the courts feel they should require evidence capable of satisfying exacting demands before being persuaded of similarity.[9]

It would therefore be hoped that scientific evidentiary techniques would be applied in this field of law which exists for the promotion of the arts and sciences.[10] Unfortunately this is not the case.

THE EXISTING STATE OF THE LAW

The traditional test of similarity is "the spontaneous impression received by the ordinary observer from a comparative viewing of the two works. . . ." A copy is said to be that which the "average reasonable man would promptly recognize, *without* any aid, suggestion or critical analysis of others, as so nearly like the original as to lead him to conclude that it had been taken from or reproduced from the original." It is not to be determined by the "fine analysis, argument and dissection of an expert, but by *ordinary observation* . . . the two works should be considered and

---

3 Hannah More, "Florio" (1786).

4 The word "plagiarism" was first applied to literary piracy in 1 A.D.; See BALL, *op. cit. supra* note 1, at 323; and compare Rudyard Kipling's "Barrack Room Ballads":

> "When 'Omer smote 'is bloomin' lyre
> He'd 'eard men sing by land an' sea;
>
> An' what he though 'e might require
> 'e went an' took . . . the same as me".

5 For a meaningful discussion of the breadth of communication's role in today's society, see Wirth, *Consensus and Mass Communication,* 13 AMER. SOCIOL. REV. 1 (1948).

6 Emerson v. Davies, 8 Fed. Cas. 615, 619, No. 4,436 (C.C.D. Mass. 1845). One may speculate as to whether plaintiff Story felt as philosophical about it two years later when his "Commentaries" were allegedly infringed. Story v. Holcombe, 23 Fed. Cas. 171, No. 13,497 (C.C.D. Ohio 1847).

7 Note, 38 CALIF. L. REV. 332, 336 (1950).

8 See 51 L.R.A. 353, 378 (1901).

9 *See, e.g.,* Simonton v. Gordon, 297 Fed. 625 (S.D.N.Y. 1924).

10 *Cf.* Beutel, *Experimental Jurisprudence,* 51 COL. L. REV. 415 (1951).

tested, not hypercritically or with meticulous scrutiny, but by the observations and impressions of the average reader and spectator." In short, the present test of similarity insists that the "common knowledge of the average reader, observer, spectator or listener is the standard of judgment which must be used."[11]

This test of the "impression upon the ordinary observer" without any aid from experts has seemingly become well settled law applicable to every kind of copyright infringement action: music;[12] motion pictures;[13] plays;[14] radio programs;[15] and, of course, all books, poems and other literary works.[16] It is especially well established, in principle at least, in the Second and Ninth Federal Circuits which, by virtue of their coverage of New York City and Hollywood, are the most important circuits in copyright law.[17] The roots of the rule may be traced back to leading copyright cases.[18]

This rule of the ordinary observer appears at first blush to be another phase of the typical legal maxim about the "ordinary reasonable man."[19] However, its more profound effect lies in its condemnation of expert opinion. Treatment of any type of expert evidence upon the issue of similarity varies from case to case, but it is consistently negative.[20]

---

[11] BALL, *op. cit. supra* note 1, at 339, 345; 20th Century Fox v. Stonesifer, 140 F.2d 579, 582 (9th Cir. 1944); Fleischer Studios v. Freundlich, 73 F.2d 276, 278 (2d Cir. 1934); King Features Syndicate v. Fleischer, 299 Fed. 533, 535 (2d Cir. 1924); Contemporary Arts v. Woolworth Co., 93 F. Supp. 739, 743 (D. Mass. 1950); Seltzer v. Sunbrock, 22 F. Supp. 621 (S.D. Cal. 1938); Hirsch v. Paramount Pictures, 17 F. Supp. 816, 818 (S.D. Cal. 1937); Echevarria v. Warner Bros., 12 F. Supp. 632 (S.D. Cal. 1935); Wiren v. Shubert Theater Corp., 5 F. Supp. 358, 362 (S.D.N.Y. 1933); Barbadillo v. Goldwyn, 42 F.2d 881 (S.D. Cal. 1930); Fox, *Evidence of Plagiarism in the Law of Copyright,* 6 TORONTO L. J. 414, 458 (1946); Miller, *A Re-examination of Literary Piracy,* 5 NEWARK L. REV. 327, 342 (1940); Note, 38 CALIF. L. REV. 332 (1950).

[12] Arnstein v. Broadcast Music, 46 F. Supp. 379 (S.D.N.Y. 1942); Hein v. Harris, 175 Fed. 875 (C.C.S.D.N.Y. 1910).

[13] Harold Lloyd Corp. v. Witwer, 65 F.2d 1, 18 (9th Cir. 1933); Solomon v. RKO Radio Pictures, 44 F. Supp. 780, 782 (S.D.N.Y. 1942); Roe-Lawton v. Hal Roach Studios, 18 F.2d 126, 128 (S.D. Cal. 1927). See note 11 *supra.*

[14] Dymow v. Bolton, 11 F.2d 690, 692 (2d Cir. 1926); Wiren v. Schubert Theatre Corp., 5 F. Supp. 358 (S.D.N.Y. 1933); BALL, *op cit. supra* note 1, at 345.

[15] Stanley v. C.B.S., 208 P.2d 9 (1949), *aff'd,* 35 Cal. 2d 779, 221 P.2d 73 (1950); BALL, *op. cit. supra* note 1.

[16] BALL, *op. cit. supra* note 1, at 321, 324, 339, 345; and Fox, *supra* note 11, at 458; see note 11 *supra.*

[17] See notes 11 to 16 *supra.*

[18] Daly v. Palmer, 6 Fed. Cas. 1132, 1138, No. 3,552 (C.C.S.D.N.Y. 1868); Dymow v. Bolton, 11 F.2d 690, 692 (2d Cir. 1926).

[19] It has been so analyzed and attributed; see BALL, *op. cit. supra* note 1, at 345; and Sheldon v. Metro Goldwyn Mayer, 7 F. Supp. 837 (S.D.N.Y. 1934).

[20] AMDUR, COPYRIGHT LAW AND PRACTICE 727 (1936).

Some courts have held that such expert testimony "ought not to be allowed at all."[21] Other courts hold that such evidence is admissible and perhaps even helpful to bring alleged similarities to the court's attention; but that such evidence has "no probative force" because it is in reality "only pleadings" or "secondary evidence"; consequently the trial court must not accept the opinions, arguments or conclusions of such experts.[22] Other courts emphasize that such testimony is completely unnecessary, since the court, assuming the role of the ordinary observer, has before it all necessary data when it has compared the two works on its own.[23]

It is not clear which chicken or egg first led to the other: the expert being rejected because the test is simply that of the ordinary observer; or the test of the ordinary observer being established as the result of a suspicion and disdain for expertise. Today the two go hand in hand. Adherence to the test of the ordinary observer is said to spring "from the fear that by dissection, abstraction and analysis of the works, similarity will be found in the wholly dissimilar."[24] Adherence to the rule against experts is said to be based on the test of the ordinary observer.[25]

What is the basis for this fear and rejection of the aid of an expert in finding similarity? A leading precedent relied upon for this point of view is the oft-quoted dictum of Judge Learned Hand:

> We cannot approve the length of the record, which was due chiefly to the use of expert witnesses. . . . The testimony of an expert upon such issues, especially his cross-examination, greatly extends the trial. It ought not be allowed at all; and while its admission is not a ground for reversal, it encumbers the case and tends to confusion, for the more the court is led into the intricacies of dramatic craftsmanship, the less likely it is to stand upon *the firmer, if more naive, ground* of its considered impressions upon its own perusal. We hope that in this class of cases such evidence may in the future be entirely excluded, and the case confined to the actual issues. . . .[26]

Again and again the courts repeat their scorn for "dissection" under

---

21 Nichols v. Universal Pictures, 45 F.2d 119, 122, 123 (2d Cir. 1930).

22 West Publishing Co. v. Edward Thompson Co., 169 Fed. 833, 854, 858 (C.C.E.D.N.Y. 1909); Encyclopedia Britannica Co. v. American Newspaper Ass'n, 130 Fed. 460 (C.C.D.N.J. 1904), *aff'd,* Werner Co. v. Encyclopedia Britannica Co., 134 Fed. 831 (3d Cir. 1905); Lawrence v. Dana, 15 Fed. Cas. 26, No. 8,136 (C.C.D. Mass. 1869); Simonton v. Gordon, 297 Fed. 625 (S.D.N.Y. 1924); AMDUR, *op. cit. supra* note 20, at 1075; BALL, *op. cit. supra* note 1, at 598, 599-601; WEIL, AMERICAN COPYRIGHT LAW 458 (1917).

23 Christianson v. West Publishing Co., 149 F.2d 202 (9th Cir. 1945); Kustoff v. Chaplin, 120 F.2d 551, 559, 561 (9th Cir. 1941); Falk v. Donaldson, 57 Fed. 32 (C.C.S.D.N.Y. 1893); BALL, *op. cit. supra* note 1, at 598.

24 Note, 38 CALIF. L. REV. 332, 333 (1950).

25 See note 23 *supra.*

26 Nichols v. Universal Pictures, 45 F.2d 119, 122, 123 (2d Cir. 1930).

the microscopic eye and cold, fine analysis and argument of the expert, which they are certain would reflect technical ingenuity rather than reality.[27] The court is fearful of being led into the intricacies of literary technique by experts who reduce incidents to abstractions or sublimations in order to find identity, and consequently prevent the judge from placing himself in the attitude of the ordinary observer.[28] The courts feel that copyrighted works (as well as the copyright law) are written for impression "upon the great multitude of plain people" and not for the few critics and experts skilled in the arts.[29] Thus if there is any literary piracy, the "ordinary person . . . should detect that fact without any aid or suggestion or critical analysis. The reaction of the public to the matter should be spontaneous and immediate."[30] Since the issue of infringement or noninfringement is thus treated on a clear and superficial basis, expert testimony, which could only confuse the mechanics of this standard of judgment, is rejected.[31] Serious efforts at comparative analysis are criticized as inapplicable and naive.[32] The courts have instead established personal impression as the ultimate, and best, test of literary similarity.[33]

---

[27] Cain v. Universal Pictures Co., 47 F. Supp. 1013, 1015-1016 (S.D. Cal. 1942); Christie v. Harris, 47 F. Supp. 39, 41 (S.D.N.Y. 1942); Carew v. RKO Radio Pictures, 43 F. Supp. 199 (S.D. Cal. 1942); Eisman v. Samuel Goldwyn, Inc., 23 F. Supp. 519 (S.D.N.Y. 1938); Frankel v. Irwin, 34 F.2d 142, 144 (S.D.N.Y. 1918); Fox, *supra* note 11, at 457-58; Miller, *A Re-examination of Literary Piracy*, 5 NEWARK L. REV. 327 (1940).

[28] See footnote 27 *supra*; see BALL, *op. cit. supra* note 1, at 599, and Fox, *supra* note 11, at 457, and cases cited therein.

[29] Arnstein v. Porter, 154 F.2d 464, 468, 471 (2d Cir. 1946); 20th Century Fox v. Stonesifer, 140 F.2d 579, 582 (9th Cir. 1944); Kustoff v. Chaplin, 120 F.2d 551, 559-561 (9th Cir. 1941); Dymow v. Bolton, 11 F.2d 690, 692 (2d Cir. 1926); Arnstein v. Broadcast Music, 46 F. Supp. 379 (S.D.N.Y. 1942); and WEIL, *op. cit. supra* note 22, at 458. But *compare* BALL, *op. cit. supra* note 1, at 344, *with* Arnold Bennett's well known observation:

> "The large majority of our fellow-citizens care as much about literature as they care about aeroplanes or the program of the Legislature. They do not ignore it; they are not quite indifferent to it. But their interest in it is faint and perfunctory; or, if their interest happens to be violent, it is spasmodic. . . . A classic is a work which gives pleasure to the minority which is intensely and permanently interested in literature."

BENNETT, LITERARY TASTE, HOW TO FORM IT (1909) in a chapter called "Why a Classic is a Classic."

[30] Harold Lloyd Corp. v. Witwer, 65 F.2d 1 (9th Cir. 1933).

[31] Burns v. 20th Century Fox Film Corp. 75 F. Supp. 986, 992 (D. Mass. 1948); ("Miracle on 34th Street" held to be no infringement of Ralph Burns' novel "Angel on Horseback"); *cf.* note 23 *supra*.

[32] See generally footnotes 11 to 32 *supra*; and see BALL, *op. cit. supra* note 1, at 346.

[33] See note 50 *infra*.

FAULTS AND FALLACIES OF THE TRADITIONAL TEST

Certainly much of the blame for the rejection of expert evidence for personal impression on the issue of similarity must be placed on the unhappy experiences courts have had with such evidence in the past. In many cases, the court has been absolutely correct in analyzing the parallel columns and elaborate diagrams submitted by experts as "illustrative of the classic difficulty of not being able to see the forest for the trees . . . not only unpersuasive but in many parts silly . . . inconsequential similarities . . . fantastic hypotheses . . . the height of absurdity . . . solemn nonsense . . . adjectives so general as to be quite useless. . . ."[34]

The judicial standing of the traditional test, with its disdain of "expertise," is not completely secure. The United States Supreme Court has never approved the test for its own use.[35] Moreover, many of the early cases relied upon as precedent were patent infringement cases; here the test of the ordinary observer is used on patent designs because the essence of the protection lies less in the claim to authorship and originality (as in copyright protection) as it does in prohibiting a resemblance sufficient to deceive purchasers of a particular article (a finished product).[36] Such a rationale is not applicable to a large majority of copyright cases; no purchaser was deceived into believing that X's film was Y's book; or even that he was seeing Z film when he actually intended to purchase tickets to Y film.

Moreover, the standard procedure in the early common law cases was to refer both works to a master for expert examination and comparison. The case of *Lawrence v. Dana*,[37] often relied upon as a precedent for the rejection of expert testimony, insisted that a comparison by the judges alone, without the master's report, was a procedure impractical in its operation and dangerous in its potential injustice to the authors. Early English and Canadian Courts also followed this procedure: "It would be absurd for the Chief Justice to sit and hear both books read over. . . . This I think is one of those cases where it would be much better for the parties to fix upon two persons of learning . . . who would ac-

[34] Frankel v. Irwin, 34 F.2d 142, 144 (S.D.N.Y. 1918); Nichols v. Universal Pictures, 45 F.2d 119, 122, 123 (2d Cir. 1930); Bachman v. Belasco, 224 Fed. 817 (2d Cir. 1915); Lewys v. O'Neill, 49 F.2d 603, 611 (S.D.N.Y. 1931); and Fox, *supra* note 11, at 449.

[35] *Cf.* White-Smith Music Publishing Co. v. Apollo Co., 209 U.S. 1 (1908).

[36] Gorham Co. v. White, 81 U.S. 511 (1842); Ripley v. Elson Glass Co., 49 Fed. 927, 930 (C.C.E.D. Ohio 1892); *cf.* Boosey v. Whight, 1 Ch. 122 (1900).

[37] 15 Fed. Cas. 26, No. 8, 136 (C.C.D. Mass. 1869); See also Greene v. Bishop, 10 Fed. Cas. 1128, No. 5,763 (C.C.D. Mass. 1858).

curately and carefully compare them, and report their opinion to the court."[38]

Moreover, the test is very often ignored by the courts, and has been specifically criticized or rejected by judges and legal writers cited herein. Nathan Burkan, a well known copyright lawyer, was at one time successful in persuading a rejection of the ordinary observer test which had been strongly urged upon the Judge by opposing counsel. Burkan convinced the judge that he must "have a more Olympian viewpoint than the average play-goer."[39] Even the Second Circuit has criticized the test as "impractical, ignored, artificial and disappointingly inaccurate in its application, and acknowledged as inconclusive."[40] In a recent musical copyright case, speaking through Judge Frank, the Second Circuit distinguished the issue of *copying*, where expert analysis and even "dissection" were now supposedly invited, from the issue of *illicit copying*, where the ordinary observer test remains.[41] In concurring, Judge Clark attacked the distinction, and concluded that copying is one "issue to be decided with all the intelligence, musical as well as legal, we can bring to bear upon it."[42] Thus expert testimonials are actually often heard and relied upon in Second Circuit copyright cases.[43]

Similarly, West Coast courts have relied upon expert testimony to show value lost, damages, orginality and other issues closely related to similarity.[44]

The same authors on copyright law who uphold the traditional test also realize its fallacious and artificial basis, and its false oversimplification of the problem:

---

38 Giles v. Wilcox, 2 Atk. 141, 144 (Ch. 1740); See also Jeffrey v. Bowles, [1770] Dick. 429; Jeffrey v. Leadbetter, 4 Ves. 681 (1799).

39 Sheldon v. Metro Goldwyn Mayer, 7 F. Supp. 837 (S.D.N.Y. 1934).

40 Shipman v. RKO Radio Pictures, 100 F.2d 533 (2d Cir. 1938); *cf.* the concurring opinion by Judge L. Hand saying the test is "necessarily vague"; *id.* at 538; and *cf.* Dellar v. Goldwyn, 104 F.2d 661 (2d Cir. 1939), saying that the above case meant no change in the rule.

41 Heim v. Universal Pictures Co., 154 F.2d 480, 488 (2d Cir. 1946).

42 *Id.* at 488, 491.

43 It might be pointed out that this is especially true in music cases; see Wilkie v. Santly Bros., 91 F.2d 978, 979 (2d Cir. 1937); Marks v. Leo Feist, 290 Fed. 595 (2d Cir. 1923); Arnstein v. A.S.C.A.P., 29 F. Supp. 388 (S.D.N.Y. 1939). And see also a much earlier Second Circuit case where expert evidence of parallel columns was highly valued by the court; West Publishing Co v. Lawyers Cooperative Publishing Co., 79 Fed. 756 (2d Cir. 1897).

44 Universal Pictures Co. v. Harold Lloyd Corp., 162 F.2d 354, 361, 370 (9th Cir. 1947); Stanley v. C.B.S., 208 P.2d 9 (1949), *aff'd*, 35 Cal. 2d 779, 221 P.2d 73, 91 (1950) (both majority and dissenting opinions criticized the oversimplification of the ordinary observer test); Golding v. RKO Pictures, 193 P.2d 153 (1948), *aff'd*, 208 P.2d 1 (1949), *aff'd*, 35 Cal. 2d 690, 221 P.2d 95 (1950).

> If it were strictly applied, substantial injuries would often result and many willful plagiarists escape the penalty of their acts. The arts are too complex to be reduced to such a simplified formula. More often than not, it requires a searching analysis based upon the testimony of expert witnesses before a charge of infringement can be either established or dismissed.[45]

To distinguish the ideas, plots, title, phraseology, characters and locale, all of which are not infringible, from the "original form of expression, language or thought sequence and literary style" is simply too difficult a job for the ordinary observer making a superficial comparison.[46] The fact that the two works produce the same emotions upon the observer is not enough,[47] and, besides, is impossible of verification. It is too much to expect the average individual to have sufficient understanding of the particular subject-matter involved[48] or to view it with a perspective common to all in order that all "ordinary observers" would report the same thing (which is surely the first requirement of *reliability* of any descriptive technique).

It has been said that many judges, in attempting to apply the test, have become lost "in a metaphysical discussion of the similarities or lack of similarities . . . and lose sight of the real problem."[49] Yet the courts now seem to consider themselves the best arbiters of originality and plagiarism, and base their conclusions on their own literary, dramatic, musical or artistic perceptions and preconceptions.[50] There surely is no logical justification for such a standard. Judges and authors agree that the former's qualifications as judges do not give them the necessary background for such a task.[51] Meanwhile, copyright counsel have no way of knowing whether the judge read all of the works, scanned part of

---

[45] Fox, *supra* note 11 at 417; see also Ball, *op. cit. supra* note 1, at 347.

[46] Ball, *op. cit. supra* note 1, at 321, 342, 347; Amdur, *op. cit. supra* note 20, at 696.

[47] Ball, *op. cit. supra* note 1, at 344; Amdur, *op. cit. supra* note 20, at 712.

[48] See note 23, *supra*.

[49] Driscoll, *Copyright Infringement*, 11 Ford. L. Rev. 63 (1942).

[50] Solomon v. RKO Radio Pictures, 44 F. Supp. 780, 782 (S.D.N.Y. 1942); Arnstein v. Marks Music Corp., 11 F. Supp. 535 (S.D.N.Y. 1935); Lowenfels v. Nathan, 2 F. Supp. 73 (S.D.N.Y. 1932); Underhill v. Belasco, 254 Fed. 838 (S.D.N.Y. 1918); "While I do not pretend to any dramatic knowledge . . ."; Fox, *supra* note 11, at 451, 457; and see generally notes 11 to 32 *supra*. "Judge Yankwich, in a recent paper, admits that the determination of similarity or originality is difficult, and that the traditional test, which he supports, requires an understanding of the medium with which the judge is dealing and a 'proper and intelligent' application of the test. But he strongly prefers it to expert evidence: 'Fortunately, American judges trained in the pragmatic system of the common law have . . . kept their feet solidly on the ground, much to the dismay of disgruntled authors . . . who . . . would substitute unreal for the pragmatic tests of originality.'" Yankwich, *Originality in the Law of Intellectual Property*, 11 F.R.D. 457, 468, 471 (1951).

[51] Hein v. Harris, 175 Fed. 875 (C.C.S.D.N.Y. 1910); (Judge L. Hand says that he is glad a local judge, who is also a musician, agrees); Miller, *supra* note 27, at 327.

each, used his own sampling system, or how much weight, if any, he will give to the exhibits submitted.[52]

In short, the traditionally popular rule of similarity is neither an established legal tradition nor unanimously popular.. Its inherent weaknesses and abuses have been set forth frequently; and its employment by most courts is an abandonment of serious effort to protect the rights of all. The only reason for retaining such a practice in the copyright courts is the lack of a better alternative and a fear of the results if the gates are opened to expert evidence. But mere criticism or dissatisfaction with the present test without specific proposals of alternatives makes no contribution toward alleviation of this dilemma and only further entrenches the traditional position.[53]

CONTENT ANALYSIS: A PROPOSED ALTERNATIVE TO THE TRADITIONAL TEST

Is it possible for expert effort to be utilized in a more scientific determination of similarity? If expert help, whether sought by court or by counsel, is to consist of "fantastic hypotheses" and "solemn nonsense," then there is no need for abandoning the existing rules. If, on the other hand, sound techniques of analysis could be employed by qualified experts, the law of copyright can progress to more adequate standards, and victory or defeat in an infringement action will be less a matter of chance.

The social sciences have developed evidentiary techniques used by literary historians[54] to a point where they have already been accepted in several phases of law. These techniques today are commonly grouped under the label of a specialized science, "content analysis."

Content analysis is a precise research technique for the objective, systematic, and quantitative as well as qualitative description of the contents of any sort of communication: newspaper stories,[55] editorials,[56]

---

[52] See Fox, *supra* note 11, at 451, 457.

[53] *See, e.g.,* Nimmer, *Inroads on Copyright Protection,* 64 HARV. L. REV. 1125 (1952), wherein the author questions the legal basis of the traditional test, fears its effect upon honest artists and authors, but opposes the exclusive use of any 'scientific' test, of which he has none to suggest, though scornfully mentioning the algebraic formulae of MALEVINSKY, THE SCIENCE OF PLAYWRITING (1925).

[54] BERELSON & LAZARSFELD, THE ANALYSIS OF COMMUNICATION CONTENT 15, 48, 49 (Preliminary draft 1948).

[55] THOMPSON, QUANTITATIVE ANALYSIS OF NEWSPAPER OPINION PRIOR TO THE SPANISH-AMERICAN WAR (unpublished master's thesis, U. of Chicago 1946); LEITES, THE THIRD INTERNATIONAL ON ITS CHANGES OF POLICY: A STUDY OF POLITICAL COMMUNICATION (Library of Congress, Experimental Division for Study of War Time Communications, Document No. 25, 1942); Lasswell, *The Politically Significant Content of the Press:*

short stories and novels,[57] radio and television broadcasts,[58] moving pictures,[59] newsreels,[60] public speeches and sermons,[61] textbooks,[62] advertising,[63] and even those non-verbal forms of communication not considered here but which are often involved in copyright cases—music, art and maps.[64] The technique is to dissect the mass of content, the whole of which might be beyond the careful perusal of any judge or jury, through rules of analysis that establish categories making possible objective selection and classification of the significant material. The use of content analysis thus enables the investigator, among other things, to determine relationships between a given communication and some other content and in more precise terms than is provided by impressionistic "more or less" judgments.[65]

Conclusions reached by the methods of content analysis have been admitted in evidence by federal courts, relied upon by the Department of Justice, and successfully applied by and before the Federal Communications Commission.[66] These uses have been chiefly in connection

Coding Procedures, 19 JOURNALISM Q. 12 (1942); Lasswell, *The World Attention Survey,* 5 PUB. OP. Q. 456 (1941); Wilson, *Newspaper Opinion and Crime in Boston,* 29 J. CRIM. L. & CRIMINOLOGY 202 (1938).

[56] Ames, *Editorial Treatment of Lynchings,* 2 PUB. OP. Q. 77 (1938); Geller, Kaplan, and Lasswell, *An Experimental Comparison of Four Ways of Coding Editorial Content,* 19 JOURNALISM Q. 362 (1942); Russell and Wright, *National Attitudes in the Far Eastern Controversy,* 27 AM. POL. SCI. REV. 555 (1933); Twohey, *An Analysis of Newspaper Opinion on War Issues,* 5 PUB. OP. Q. 448 (1941).

[57] McKenzie, *Treatment of War Themes in Magazine Fiction,* 5 PUB. OP. Q. 227 (1941).

[58] Lasswell, *Describing the Contents of Communications,* in SMITH, LASSWELL, AND CASEY, PROPAGANDA, COMMUNICATION, AND PUBLIC OPINION (1946).

[59] Jones, *Quantitative Analysis of Motion Picture Content,* 6 PUB. OP. Q. 411 (1942).

[60] Dale, *Need for Study of the Newsreels,* 1 PUB. OP. Q. 122 (1938).

[61] Hamilton, *Social Optimism and Pessimism in American Protestantism,* 6 PUB. OP. Q. 280 (1942); McDiarmid, *Presidential Inaugural Addresses: A Study in Verbal Symbols,* 1 PUB. OP. Q. 79 (1937).

[62] Saunders, *Social Ideas in the McGuffey Readers,* 5 PUB. OP. Q. 579 (1941).

[63] Severson, *Nationality and Religious Preferences as Reflected in Newspaper Advertisements,* 44 AMER. J. SOCIOL. 540 (1939); Shuman, *Identification Elements of Advertising Slogans,* 17 S.W. SOC. SCI. Q. 342 (1937).

[64] BERELSON & LAZARSFELD, *op. cit. supra* note 54.

[65] Note *Content Analysis: A New Evidentiary Technique,* 15 U. OF CHI. L. REV. 910, 912, 914, 915 (1948); SMITH, LASSWELL, & CASEY, PROPAGANDA, COMMUNICATION AND PUBLIC OPINION (1946); Janis, *The Problem of Validating Content Analysis,* and Kaplan and Goldsen, *The Reliability of Content Analysis Categories,* in LASSWELL, LEITES *et al.,* LANGUAGE OF POLITICS, 55 *et seq.,* 83 (1949).

[66] See Note, 15 U. OF CHI. L. REV. 910 (1948); Lasswell, *Detection: Propaganda Detection and the Courts* in LASSWELL *et al., op. cit. supra* note 65, at 173; U.S. v. Pelley, 132 F.2d 170 (7th Cir. 1942), *cert. denied,* 318 U.S. 764; and see U.S. v. Auhagen, 39 F. Supp. 590 (D.D.C. 1941); see *In re* W.H.K.C., 10 F.C.C. 515 (1945); see United States v. German-American Vocational League, 153 F.2d 860 (3d Cir. 1945), *cert. denied,* 328 U.S. 833 (1946); United States v. Transocean, Docket No. 67418 (D.C. 1941).

with the detection of enemy propaganda and sedition, in investigating charges of bias, and in comparing radio broadcasting content with administrative standards. Its successful use in these fields suggests the possibility of utilizing the technique elsewhere in the law, bringing to courts "facts heretofore beyond their ken."[67] The courts have already recognized such evidence to be competent expert testimony which satisfies the guarantee of trustworthiness for criminal action; and not within the hearsay rule since it goes to proving the fact of utterance, not the truth of the utterance.[68]

The field of copyright law and proof of similarity would seem especially adaptable to modern scientific methods of content analysis. The latter's use of the "parallel test" in analyzing similar themes and differences, repeated or copied errors, sources, distinctive vocabularies and emphasis in presentation[69] should be useful in providing courts and counsel with more adequate methods of detection in difficult infringement actions. If the courts wish to cling to the shell of the "ordinary observer" test, copyright counsel in the future can introduce scientific evidence as to what the actual impressions on ordinary observers are through reliable methods of content analysis (in fact defined as describing "the influence upon typical readers"),[70] rather than through the impressions of judge or jury assuming to speak for all "ordinary observers."

A complete description and appraisal of the scientific development of content analysis is outside the scope of this paper. The succeeding paragraphs, however, will discuss some of the problems to be anticipated in utilizing this technique in the courts.

The use of content analysis techniques in literary infringement cases would be strictly limited to problems of proof regarding *content similarity*. No verifiable conclusions can be drawn regarding the intent of the writer by an analysis of what he has written.[71] Content analysis, even should

[67] See note, 15 U. OF CHI. L. REV. 910, 918, 924, 925 (1948); and Reisman, *Democracy and Defamation,* 42 COL. L. REV. 1282, 1307 (1942).

[68] See U.S. v. Pelley, *supra* note 66 at 176, 178, 180; and Brief for the U.S., 49-51; see ST. GEORGE & DENNIS, A TRIAL ON TRIAL (1946), a book written by two of the leading defense counsel in the mass sedition trial where they attack content analysis as "learned hokum" (p. 302).

[69] Lasswell, *op. cit. supra* note 65, at 180.

[70] WAPLES, BERELSON AND BRADSHAW, WHAT READING DOES TO PEOPLE 146 (1940); see also Note, 15 U. OF CHI. L. REV. 910, 922 (1948).

[71] "As to the publisher's state of mind no demonstration has been made of which of the aspects of the crime story, jointly or severally, influenced the city editor in deciding to run the story: The Jewish "Angle," the sensational details, or the opportunity to spread many pictures of the murdered beauty on the front page." Note, 15 U. OF CHI. L. REV. 910, 919 (1948), regarding efforts of the American Jewish Congress to prove an anti-

it reveal a word for word similarity between two documents, is not designed to prove or disprove that the author of one document intended to copy the words of another. Content analysis does, however, make it possible to isolate, classify, and inventory quantitatively the words and themes—thus presenting a statistical opportunity to figure the probabilities that such similarity did not occur by chance.

Nor can content analysis techniques be utilized to demonstrate infringement because similar reader reactions are found to take place with respect to compared content. Who can prove the causal relationship between allegedly similar content and similar emotional reactions regardless of their correlation? To assert that infringement is proved because identical emotions do or may be predicted to occur in response to the contents of two documents under observation is to foreclose consideration of the nature of the receiving mechanisms (the psychological makeup of those to whom the materials are communicated) and the social and political climate in which the materials are received. While studies of this sort are extremely important to the field of communications,[72] they deserve no consideration by the courts because of their complexity and the huge amount of guess work still associated with the results.[73]

The two major types of content analysis which concern us have been referred to as (1) "sign-vehicle" analysis and (2) "semantical content" analysis.[74] The simpler of the two is the sign-vehicle analysis type. Its function would be to count the number of instances in which given verbal symbols or combinations of verbal symbols (words, phrases and sentences) appear. Providing that the analyst is personally dependable, fully understands the rules of analysis and what he is looking for, and is possessed of adequate qualities of perception, a first hand census of the frequency of physical occurrences is all that is involved. For many years, parallel charts attempting this type of analysis have been offered in copyright cases, generally to be rejected.

---

semitic approach of a newspaper's news columns. WBNX Broadcasting Co. *et al.*, Docket No. 6013 and *In re* Application of News Syndicate Co., Docket No. 6175, before the Federal Communications Commission.

[72] See Smith, *The Political Communication Specialist of Our Times,* and Lasswell, *Describing the Effects of Communications* in SMITH, LASSWELL, & CASEY, *op. cit. supra* note 65.

[73] The failure to achieve validity with content analysis techniques which classify language according to their probable causes and effects is discussed in Janis, *Meaning and the Study of Symbolic Behavior,* 6 PSYCHIATRY 425 (1943) and Janis, Fadner, and Janowitz, *The Reliability of a Content Analysis Technique,* 7 PUB. OP. Q. 293 (1943).

[74] These two classifications along with "Pragmatical content" analysis are discussed in Janis, *supra* note 65, at 57.

The technique of validating this content analysis type would be simple in nature. Verification would demand only the independent comparison between analyst's tabulations and frequency of occurrence of the items he was counting. Provision for an independent check of the analyst's work could be made in this manner: assume that a check was made on the analyst's tally of the number of times that the word "psychiatrist" appeared in a twelve-page short story. The analyst should divide the work under analysis into twelve or possibly twenty-four sections of equal length. His own work sheet should be divided correspondingly. Thus an independent checker would observe the results in terms of specific section and randomly spot check three or four sections in their entirety to ascertain the analyst's accuracy. If there is any deviation between the two, analyst and checker would re-examine their tallies because no deviation should exist. If a discrepancy still remains, the matter should be submitted to three referees who would remove the issue from dispute either by discovering an error on the part of one or both of the previously mentioned parties, or by clarifying the rules of analysis and investigating the orientation of analyst and checker to similar rules which should not be dissimilarly applied. This form of validation can of course be applied to any analysis technique.

It is questionable, however, that sign-vehicle analysis will offer very much utility because (a) it is not the number of times a given word or phrase may appear but the fact and circumstances of its appearance which would be important; (b) unless a literary work were a straight copy of another, the similar words or phrases seldom appear, a point which indicates the need for utilizing "semantical content" analysis techniques.

It is obvious that the fact that the word "lazy" appears ten times in short story X and ten times in short story Y (published ten years later) suggests no test of legal similarity between the two. If, on the other hand, each of the two documents had five personalities figuring in its plot, if all five characters of document X were respectively described by the identical words in 84 percent of the instances of description in document Y, and if eight of document Y's eleven plot situations were described equally similarly to those of document X, a conclusion would be in order as to whether or not the pattern of similarity was due to chance.

However, it is unlikely that there will be any one-to-one correlation between the chronological order of appearance of similar individuals and plot situations even though the identical words may be used. Therefore, the sign-vehicle analyst will be required to break down the content he is

studying into elements less basic than atomistic for the purpose of interpreting his analysis of verbal symbols. Thus, in analyzing the contents of documents X and Y, he would do a sign vehicular analysis of each personality introduced, each plot situation described, and each one of the props (physical environment) described. Treating both short stories in this manner, the analyst would then be able to square character against character, plot situation against plot situation, etc., in an effort to determine whether one or more units from the one document bears similarity (and to what degree) to one or more of the other.

Of even greater importance for the application of content analysis techniques to the issue of similarity in literary copyright cases are the semantical content analysis techniques. These procedures which would classify language symbols in terms of their meanings as well as their physical occurrence are generally considered to be of three fundamental types: (a) designations analysis, which enumerates the frequency with which certain objects, including persons, things, groups or concepts, are referred to (*e.g.*, references to psychiatry); (b) attribution analysis, which enumerates the frequency with which certain descriptions or characterizations are referred to (*e.g.*, Freudian); (c) assertions analysis, which enumerates the frequency with which certain objects are described or characterized in a particular way—the manner in which (a) and (b) are combined (*e.g.*, Psychiatry's origins are Freudian in nature).

Since meaning rather than only the physical record of the word is involved, two methodological problems deserve careful thought:

(1) A given portion of content could be ambiguous in its meaning for several individuals, each thinking it means something else. Assertions which are of confirmed ambiguity can, of course, be omitted, or if their omission would appear to be significant, a numerical weight can be assigned to such figures.[75]

These same problems have already been encountered in the use of personal documents in the social sciences.[76] An initial fundamental objection to conclusions emerging from studies of communications of a personal nature, *i.e.*, diaries, letters, autobiographies, etc., was that judges, no matter what their expert qualifications, could not agree upon what

[75] For a technical discussion of this point, see Kaplan and Goldsen, *The Reliability of Content Analysis Categories* (chapter V) in LASSWELL, LEITES *et al.*, LANGUAGE OF POLITICS (1949).

[76] For a full discussion of the problems of reliability and validity in connection with the analysis of personal documents, see ALLPORT, THE USE OF PERSONAL DOCUMENTS IN PSYCHOLOGICAL SCIENCE (1942); GOTTSCHALK, KLUCKHOHN, & ANGELL, THE USE OF PERSONAL DOCUMENTS IN HISTORY, ANTHROPOLOGY AND SOCIOLOGY (1945).

these documents actually said. Testimonials regarding the nature of personal document content, it was argued, would offer no meaningful conclusions regarding the content of the document.

Published studies have since considerably disspelled this challenge. In one study the conclusions of independently rated autobiographies were compared with the recorded outcome of clinical interviews for the degree of agreement with reference to the subjects' emotional stability. The degree of correlation obtained in all instances was .80 or above.[77] Allport summarizes the results of another study[78] as follows:

> Four judges rated 238 topical autobiographies on a graphic rating scale in respect to the autobiographer's attitude toward prohibition. By averaging the intercorrelations of the composite ratings, expressed as standard scores, of each pair of judges with each other pair, a reliability coefficient of +.96 was obtained. Correlating the composite ratings given by readers of the case history with the earlier measured attitude toward prohibition of the subject . . . a validity coefficient of +.81 resulted (+.86 if corrected for attenuation). The conclusion of this straightforward investigation is that judges are able to agree in their rating of self-written documents and, furthermore, to agree with an independent source of information concerning the attitudes of the writers of these documents.[79]

This problem of reliability, as it affects the analysis of many kinds of communication content, has been carefully worked out and published by content analysis experts and is tested by determining the degree of correlation between the results obtained when different analysts independently analyze the same content.[80]

(2) Another major aspect has to do with the relationship between one portion of the contents under observation and the contents as a whole. Content analysis must demand more than a theme-by-theme analysis of content so that the relationship between each part or the whole and its parts does not escape measurement. "By the nature and meaning of a statement is not meant its grammar or inner logic."[81] To what extent, for example, may given statements or themes of one document compare similarly to those of another document, yet function entirely differently in the document as a whole. A given idea or situation or personality characteristic may be exploited in any number of several ways by writers; surely none need have the same destiny or source. The expert analyst,

---

77 STAGNER, THE PSYCHOLOGY OF PERSONALITY (1937).

78 STOUFFER, AN EXPERIMENTAL COMPARISON OF STATISTICAL AND CASE HISTORY METHODS OF ATTITUDE RESEARCH (Unpublished dissertation, U. of Chicago 1930).

79 Allport, *op. cit. supra* note 76 at 24. See also Cavan, Hauser, and Stouffer, *Note on the Statistical Treatment of Life History Material,* 9 SOC. FORCES 200 (1930).

80 Janis, *supra* note 65, at 56 and note 84 *infra.*

81 Garber, *Propaganda Analysis—To What Ends?*, 48 AMER. J. SOCIOL. 240 (1942).

then, will recognize the need for full acquaintance with the document as a whole and will appreciate its contribution to the meaning of the content being analyzed.

Major variable factors which, unlike the present system of similarity comparison, can be controlled or accounted for under content analysis, are summarized by Berelson and Lazarsfeld as follows:

a) Formulation of definitions and rules: the more precise and complete the rules of analysis and the definitions of categories (and the fuller their illustration), the higher the reliability.

b) Units of analysis: the larger the units, the higher the reliability for subject matter categories (and perhaps for other categories as well).

c) Training and experience of the coders: the more general experience in content analysis and the more specific training on the particular project the coders have had, the higher reliability is likely to be.[82]

d) Complexity of the category: the more elaborate and complex the category system, involving many facets of the content, the lower the reliability is likely to be.[83]

As was previously mentioned, it is not the purpose of this paper to examine the qualitative and statistical tests of the reliability of content analysis techniques. The reader may check the several studies to decide for himself.[84] Suffice to say, there are many published studies whose results have been tested for reliability and found satisfactory, although universal generalizations of reliability have yet to emerge. Unfortunately some studies indicate that no effort was made to determine the reliability of their results. Unpublished studies as well as those which may have been contemplated but abandoned for reasons of want of reliability cannot be cited.

Content analysis, comparatively toddling in scientific infancy, deserves cautious employment. The potential expense of the expert testimony required suggests the danger that wealthy litigants could utilize it while impoverished litigants could not. Its novelty in the courtroom implies that some judges will encourage its use by the parties and some will not. It would seem advisable for the parties to agree in advance, in much the same way as was suggested over 200 years ago, on the use of experts

82 See Meier and Lervinski, *Occupational Variation in Judging Trends in Public Opinion*, 2 PUB. OP. Q. 422 (1938).

83 BERELSON & LAZARSFELD, *op. cit. supra* note 54, at 136.

84 See *id.* at 130-134 for a summary of the reliability tests and results appearing in the published reports of 18 content analysis projects. See also chapters 5-9 in LASSWELL, LEITES *et al., op. cit. supra* note 65; and BERELSON, CONTENT ANALYSIS IN COMMUNICATION RESEARCH (1951).

in content analysis.[85] Standards of analysis could also be stipulated since logical reasoning rather than subjective judgment could be utilized.

CONCLUSION

It is submitted that techniques of content aualysis provide both necessary and reliable aids to proof of similarity in copyright infringement cases. The present ordinary observer rule which excludes such expert evidence is open to error, dangerous and illogical in its use, and lacking in firm legal foundation. We no longer live in an age in which legalistic rules of this nature must be tenaciously preserved at the expense of scientific advancement. Adequate protection of our arts and sciences would seem to require an amending of the existing rules in order to consider more fully the contributions of experts skilled at techniques of content analysis.

[85] See note 67 *supra.*