

TRADEMARK CONFUSION REVEALED: AN EMPIRICAL ANALYSIS

DARYL LIM*

The likelihood of confusion standard defines the scope of trademark infringement. Likelihood of confusion examines whether there is a substantial risk that consumers will be confused as to the source, identity, sponsorship, or origin of the defendants' goods or services. This Article presents a contemporary empirical analysis of the various factors and how they interact. Conventional wisdom teaches us that courts should comprehensively traverse each factor and that likelihood of confusion cases generally require jury determination. However, the data reveals that neither is true. Instead, courts provide early off-ramps to litigants by "economizing," and analyzing only a handful of factors or by "folding" factors within each other. The findings also reveal (1) which forums are pro-defendant and which are pro-plaintiff; (2) the impact of rivalry and fair use on outcomes; and (3) an apparent Ninth Circuit dominance.

What constitutes "confusion" remains highly subjective and difficult to evaluate. Proxies like intent, survey evidence, mark strength, and consumer sophistication fail to incorporate real-world purchasing conditions or are better considered within omnibus factors. In contrast, actual confusion, mark similarity, and competitive proximity provide judges with a potent trio of factors to guide the infringement inquiry. Together with safe harbors for descriptive and expressive uses, these rules of thumb enable courts to resolve trademark disputes more coherently, consistently, and expeditiously. This Article concludes with a blueprint of how these rules of thumb complement artificial intelligence systems

* Professor of Law & Director, Center for Intellectual Property, Information and Privacy Law, *University of Illinois Chicago School of Law*. I thank Professor Y. Samuel Wang from the Department of Statistics and Data Sciences at *Cornell University* for his valuable advice on the statistical aspects of this Article. My sincere thanks to Margaret Smiley Chavez, Steve Fisher, Annemarie Gregoire, Sarah Hampton, Nicole Robinson, Kelly Welsh, and their colleagues at the *American University Law Review* who contributed to this Article. Their courtesy, professionalism, careful editorial work, and thoughtful comments are exemplary.

and how those systems can use empirical studies as training data to inform future likelihood of confusion analyses.

TABLE OF CONTENTS

Introduction	1286
I. Studying Confusion	1290
A. Methodology	1291
B. Blends, Triggers, and <i>Polaroid</i> Factors	1296
1. Blends and triggers	1297
2. The <i>Polaroid</i> factors	1300
C. Missing the Point on Consumer Confusion	1306
1. Intent	1306
2. Surveys.....	1312
3. Mark strength	1316
4. Consumer sophistication	1319
D. Coherence-Based Reasoning.....	1320
II. Rules of Thumb.....	1324
A. Actual Confusion.....	1324
B. Mark Similarity	1328
C. Competitive Proximity	1332
D. Summing It Up	1338
E. A Word on Fair Use.....	1340
III. Observations and Implications.....	1345
A. Factor Folding	1345
B. Early Off-Ramps	1347
C. Deploying Artificial Intelligence	1353
1. Conception	1354
2. Execution.....	1355
3. Limitations.....	1358
a. “Garbage-in, garbage-out”	1358
b. Biases	1359
c. Contextualizing the purchasing conditions.....	1362
d. Coding challenges	1364
Conclusion.....	1365

INTRODUCTION

Consumers rely on a consistent commercial lexicon to reduce mental costs associated with purchasing decisions and in turn commercial enterprises gain an incentive to invest in quality products

and services.¹ Businesses imbue words, symbols, scents, and sounds with information about their goods and services.² In turn, consumers rely on this imbued information to navigate their decisions by making snap judgments about the price and quality of products or services they purchase.³ Thus, electric vehicle enthusiasts may seek out Tesla's stylized "T" comprising a stator and rotor, and connoisseurs of Chick-fil-A's chicken sandwiches will scout for its distinctive red-on-white text.

When trademark owners assert their rights, courts apply the likelihood of confusion standard, which seeks to determine whether the defendants' use of a similar mark is likely to cause consumer confusion. The likelihood of confusion standard is the linchpin of trademark infringement.⁴

Unfortunately, what constitutes "confusion" remains highly subjective and difficult to evaluate.⁵ Additionally, the likelihood of confusion standard remains poorly theorized and judges applying the standard often fail to adequately explain their analyses in their opinions in a way that future courts can easily apply and replicate.⁶

When defendants counterfeit the trademark outright, liability is clear. Literal infringement has occurred. However, trademarks protect their owners beyond literal infringement like patents and copyrights.⁷ Nonliteral infringement can occur when, for example, there is a colorable difference in the marks. This exposes parties to uncomfortably

1. *Qualitex Co. v. Jacobson Prods. Co.*, 514 U.S. 159, 163–64 (1995) (“[T]rademark law . . . reduces the customer’s costs of shopping and making purchasing decisions, . . . for it quickly and easily assures a potential customer that this item—the item with this mark—is made by the same producer as other similarly marked items that he or she liked (or disliked) in the past.”).

2. See 1 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 3, at 1 (5th ed., 2010).

3. *Qualitex Co.*, 514 U.S. at 163–64.

4. See *infra*, Part I. The Lanham Act prohibits the use of a registered mark in a manner “likely to cause confusion,” 15 U.S.C. § 1114(1)(a), as well as the use of any term or name in a manner “likely to cause confusion” about the affiliation of the user with another person. *Id.* § 1125(a)(1); see also *Mil. Ord. of Purple Heart Serv. Found., Inc. v. Mil. Ord. of Purple Heart of U.S., Inc.*, 852 F. App’x 6, 9 (D.C. Cir. 2021).

5. See *infra* Part I.

6. See *infra* Part II; see also Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581, 1582 (2006) (“Its current condition is Babelian.”).

7. See generally Daryl Lim, *Judging Equivalents*, 36 SANTA CLARA HIGH TECH. L.J. 223 (2020) (tracing the origins of the doctrine of equivalents and explaining the rationale behind the doctrine—to protect intellectual property owners from infringers seeking to evade liability by making insubstantial changes) [hereinafter Lim, *Judging Equivalents*].

uncertain waters.⁸ Patent law has claims to give notice of a patent's metes and bound.⁹ Neither trademark nor copyright law has claims, leaving courts unclear on operationalizing technical similarity or market substitution considerations.¹⁰

In a negative feedback loop, indeterminacy over the likelihood of confusion standard muddies trademark law's focal point and scope, while polluting adjacent disciplines like copyright and patent law.¹¹ Proper notice about the existence and scope of legal rights is critical to any property system, but especially trademark rights, because trademarks last indefinitely, meaning their owners obtain a timeless monopoly without the same limitations and threshold requirements placed on patent and copyright holders.¹² A patchwork of inconsistent results destabilizes the system for everyone.¹³ Indeterminacy has many negative impacts, including causing negotiations to break down, which harms *both* brand owners and potential licensees, and acting as a drag on dispute resolution, compliance, and social equity.¹⁴ Indeterminacy also acts as a drag on dispute resolution, compliance, and social equity.¹⁵ The rational response must be a call for clarity in the law.

The likelihood of confusion standard examines whether consumers will likely be confused as to the source, identity, sponsorship, or origin of the goods and requires "a substantial likelihood that the public will

8. Michael Grynberg, *Thick Marks, Thin Marks*, 67 CASE W. RES. L. REV. 13, 15 (2016) ("Many open questions in modern trademark law concern which parts of the range belong under the trademark holder's control.").

9. 35 U.S.C. § 112(b) (requiring patentees to include in their patent "one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor . . . regards as the invention").

10. Jeanne C. Fromer & Mark A. Lemley, *The Audience in Intellectual Property Infringement*, 112 MICH. L. REV. 1251, 1296–99 (2014).

11. See Robert G. Bone, *Notice Failure and Defenses in Trademark Law*, 96 B.U. L. REV. 1245, 1255 (2016) ("[W]hat makes the scope of rights so uncertain is the vagueness of the likelihood-of-confusion test ('LOC test') for infringement.") [hereinafter Bone, *Notice Failure and Defenses*]; Amy Adler & Jeanne C. Fromer, *Taking Intellectual Property into Their Own Hands*, 107 CALIF. L. REV. 1455, 1523 (2019) ("Trademark law is similarly complex and unpredictable with regard to important doctrines.").

12. See William M. Landes & Richard A. Posner, *Trademark Law: An Economic Perspective*, 30 J.L. & ECON. 265, 287 (1987) ("The lack of a fixed term for trademarks is one of the striking differences between trademarks, on the one hand, and copyrights and patents, on the other."). See generally 15 U.S.C. § 1114 (discussing infringement and remedies).

13. Thomas H. Watson, *Pay Per Click: Keyword Advertising and the Search for Limitations of Online Trademark Infringement Liability*, 2 CASE W. RES. J.L. TECH. & INTERNET 101, 122 (2011).

14. See, e.g., Bone, *supra* note 11, at 1258.

15. Daryl Lim, *AI, IP, Algorithms, and Inequality*, SMU L. REV. (forthcoming 2022).

be confused.”¹⁶ Each circuit court has enumerated factors relevant in analyzing likelihood of confusion. Courts use proxies for consumer confusion like intent, survey evidence, mark strength, and consumer sophistication.¹⁷ However, these either fail to incorporate real-world purchasing conditions as a doctrinal matter or are better considered part of a streamlined likelihood of confusion test.¹⁸ The trio of actual confusion, mark similarity, and competitive proximity provides judges a small but potent cluster of factors.¹⁹ Together with safe harbors for descriptive and expressive uses, these enable a more coherent, consistent, and expedient resolution of trademark disputes.

This Article presents a contemporary empirical analysis of each likelihood of confusion factor and how they interact with one another. Conventional wisdom teaches us that courts should comprehensively traverse each factor and that likelihood of confusion cases generally require jury determination.²⁰ The data reveals that neither is true. Instead, courts provide early off-ramps to litigants by “economizing” and applying just a handful of factors or by “folding” factors into each other in grouped layers.²¹ The findings also reveal the Ninth Circuit’s dominance, pro-defendant and pro-plaintiff fora, and the impact of rivalry and fair use on case outcomes.²² This Article concludes with a blueprint of how artificial intelligence (AI) systems can use empirical studies as training data to help stakeholders make and predict confusion analyses.²³

Part I introduces this empirical study’s methodology, goals, and limitations before elaborating on this Article’s doctrinal and policy impetus. The discussion charts how blending technical trademarks and trade names along with the expansion of triggers ensnares defendants in trademark liability. It then shifts to make critical observations gleaned from the data, including the impact of rivalry on modern case outcomes, the dominance of the Ninth Circuit in federal trademark litigation, and the most pro-defendant and pro-plaintiff circuits today.

16. *Fisher Stoves, Inc. v. All Nighter Stove Works, Inc.*, 626 F.2d 193, 194 (1st Cir. 1980).

17. *See infra* Section I.C.

18. *See generally* 4 MCCARTHY, *supra* note 2, § 23:1 (criticizing the “amorphous and indefinite” nature of the multi-factor likelihood of confusion test for creating legal unpredictability, increasing litigation costs, and chilling socially valuable uses).

19. *See infra* Part II.

20. *See infra* Part II.

21. *See infra* Part III.

22. *See infra* Part I.

23. *See infra* Part III.

Part I also presents a detailed doctrinal and empirical analysis of prominent features in the likelihood of confusion analysis—the intent of litigating parties, consumer surveys, mark strength, and consumer sophistication—and explains why each factor leads courts tasked with ascertaining consumer confusion astray. Finally, the Part introduces coherence-based reasoning and argues that a more compact nucleus of factors would better serve courts analyzing likelihood of confusion.

Part II explains why actual confusion, mark similarity, and competitive proximity offer courts that compact troika in simplifying likelihood of confusion analysis. The data reveals how most courts rely on these three factors while either paying lip service to or completely ignoring the other factors. It also shows appellate courts are complicit in this “wink-and-nod” practice, affirming lower courts in over 80% of cases on appeal. Finally, Part II explains why this trio of factors plus the fair use safe harbors of descriptive and expressive uses should form trademark law’s rules of thumb for infringement.

Part III examines the implications of the empirical study on trademark doctrine and practice. First, it observes that while the likelihood of confusion factors may present themselves as discrete categories, the dataset reveals that courts do not regard them as such. Instead, courts combine factors and analyze them both creatively and rationally. Second, courts rely on a small number of factors to economize their decisions and give parties an early off-ramp. Third, Part III explains how the empirical analysis provides a blueprint for algorithmic adjudication using AI, taking the reader from conception to execution while identifying and addressing its limitations.

I. STUDYING CONFUSION

Over the years, the jurisprudential roots of trademark law became unruly and tangled. Unfair competition intermingled with consumer protection as the Lanham Act of 1946²⁴ (“the Act”) blended trade names and technical trademarks.²⁵ A later legislative revision untied likelihood of confusion from source confusion, and courts introduced idiosyncratic rules of affiliation and sponsorship as triggers for consumer confusion.²⁶

Within the likelihood of confusion tests, factors such as defendants’ intent, survey evidence, and consumer sophistication provided a convenient but misguided attempt to get a handle on the arduous task

24. 15 U.S.C. § 1051–1127.

25. *See infra* Section I.B.

26. *See infra* Section I.B.

of determining what trademark infringement had become.²⁷ Judges often resorted to coherence-based reasoning. Once a judge determined the satisfaction of the selected factors, the judge would decide that all the other factors were present.²⁸ Though it made their work easier, it muddled the waters for everyone else. This Part presents an empirical analysis of these issues and their implications. It begins by discussing the empirical methodology.

A. Methodology

This empirical study draws upon the well-developed method of case content analysis.²⁹ The method systematically dissects a sample of judicial opinions to record consistent features, draw inferences, and uncover trends.³⁰ This social science approach to the law complements and augments traditional legal analysis.³¹ As a testament to its outsized contribution to the literature, case content analysis generates an average of seventy-seven citations per article in a milieu where 40% of law review articles receive no citations at all.³²

The value of case content analysis lies in uncovering patterns in judging. Judges routinely rely on a remarkably limited number of factors in reaching their conclusions.³³ Instead, they employ

27. See *infra* Section I.B.

28. See *infra* Section I.D.

29. Mark A. Hall & Ronald F. Wright, *Systematic Content Analysis of Judicial Opinions*, 96 CALIF. L. REV. 63, 73 (2008) (“Content analysis has proven useful for studying a broad range of legal subject areas.”); *id.* (listing “areas as far-ranging as administrative law, constitutional law, corporate and securities law, criminal law and procedure, contracts, employment discrimination, health law, and torts”).

30. See KLAUS KRIPPENDORFF, *CONTENT ANALYSIS: AN INTRODUCTION TO ITS METHODOLOGY* 18 (2d ed., 2004) (defining content analysis as “a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use”).

31. Hall & Wright, *supra* note 29, at 74; see also *id.* at 65 (“The method also helps a researcher to sort out the interaction of multiple factors that bear on an outcome in the legal system.”); *id.* at 78 (“Its strength is to provide an objective understanding of a large number of decisions where each decision has roughly the same value.”); Alan L. Tyree, *Fact Content Analysis of Case Law: Methods and Limitations*, 22 JURIMETRICS J. 1, 23 (1981) (explaining that these methods have “considerable power for the discovery of anomalies which may escape the naked eye”).

32. Hall & Wright, *supra* note 29, at 74 (“[C]ontent analysis projects appear somewhat more likely to generate discussion and citation than law review articles more generally.”).

33. See, e.g., Beebe, *supra* note 6, at 1602 (“The data collected for this study support the general hypothesis that decision makers, even when making complex decisions, consider only a small number of factors and the more specific hypothesis that, in doing so, decision makers use a core attributes heuristic.”).

heuristics—such as the likelihood of confusion factors—to cut through what would otherwise be a morass of information that could paralyze decision-making entirely.³⁴ This approach, however, makes it difficult to draw broader conclusions to inform future cases coherently.³⁵ Scholars employ case content analysis to parse through court decisions and study how judges and juries apply rules to facts in practice to address this limitation.³⁶ Thus, while the interpretive method evaluates legal principles, case content analysis “combines the analytical skills of the lawyer with the power of science that comes from articulated and replicable methods of reading and counting cases.”³⁷ In so doing, case content analysis yields useful information that moves the discussion toward a greater understanding of the bigger policy questions and helps uncover areas for further research.³⁸ It also avoids selection bias issues, which hamper the representativeness of other methods.³⁹

This Article relied on an expansive pool of 188 cases covering nearly 5,000 datapoints based on a Westlaw search for all trademark infringement cases discussing likelihood of confusion over five years between September 30, 2016, and October 1, 2021.⁴⁰ The study

34. See, e.g., Daryl Lim, *Retooling the Patent-Antitrust Intersection: Insights from Behavioral Economics*, 69 BAYLOR L. REV. 124 (2017).

35. See, e.g., Daryl Lim, *The (Unnoticed) Revitalization of the Doctrine of Equivalents*, ST. JOHN'S L. REV. (forthcoming 2022) [hereinafter Lim, *Doctrine of Equivalents*].

36. See, e.g., Lee Petherbridge, *On the Decline of the Doctrine of Equivalents*, 31 CARDOZO L. REV. 1371, 1380 (2010) [hereinafter Petherbridge, *Decline of the Doctrine of Equivalents*] (“Content analysis is capable of helping scholars verify, analyze, or refute empirical claims about case law, and it is to that purpose the approach is put in this study.”). For earlier studies where I employed a similar methodology, see Lee Petherbridge, Jason Rantanen, & Ali Majibi, *The Federal Circuit and Inequitable Conduct: An Empirical Assessment*, 84 S. CALIF. L. REV. 1293, 1304 (2011); Hall & Wright, *supra* note 29, at 77; DARYL LIM, PATENT MISUSE & ANTITRUST: EMPIRICAL, DOCTRINAL & POLICY PERSPECTIVES 8–9 (2013); Lim, *Judging Equivalents*, *supra* note 7.

37. Hall & Wright, *supra* note 29, at 100.

38. Karen A. Jordan, *Empirical Studies of Judicial Decisions Serve an Important Role in the Cumulative Process of Policy Making*, 31 IND. L. REV. 81, 88 (1998); see also Lon L. Fuller, *An Afterward: Science and the Judicial Process*, 79 HARV. L. REV. 1604, 1622 (1965) (“[P]ossible gain from researches of this kind lies in the realm of serendipity. A puzzling correlation that violates normal anticipations may set our minds going along new paths and yield unexpected insights.”).

39. Hall & Wright, *supra* note 29, at 102 (“All of these were universal samples restricted only by date, subject matter, jurisdiction, and/or source. In short, empirical researchers studying case law are usually able to avoid the selection bias issues that plague most other areas of social science.”).

40. For an example of another recent empirical study on trademarks that starting and ending during the calendar year, see Lisa Larrimore Ouellette, *The Google Shortcut to Trademark Law*, 102 CALIF. L. REV. 351, 373 (2014) (“New U.S. federal court decisions

omitted thirty-two cases captured by Westlaw's search results because they did not specifically discuss the likelihood of confusion factors.⁴¹ For accuracy, the dataset distinguishes between procedural wins (for instance, defeating a motion for summary judgment) and substantive wins on the merits (which result in a finding of infringement or non-infringement).⁴² This Article initially used Excel to hand-code the data before using IBM SPSS Statistics 28.0 to generate the graphs and crosstabs data.⁴³

The dataset of hand-coded cases included variables such as (1) the decision's date; (2) the judicial circuit; (3) whether a district or appellate court decided the case; (4) the parties' relationship as rivals (or not); (5) the procedural posture; (6) the type of mark; (7) the test employed; (8) whether the opinion discussed survey evidence; (9) which party a likelihood of confusion factor favored; (10) whether courts "folded" factors together; (11) case outcomes; and (12) whether the court discussed fair use.⁴⁴

Like all empirical studies, this one has its caveats. There are several well-recognized limitations to case content study databases. First, coding may result in incomplete or inaccurate coding, despite cross-coding and verification using a population sample.⁴⁵ Given that the focus is on features of written decisions, the data remains valid as long as it is recognized to refer to a specific population rather than a sample of all cases in all possible worlds.⁴⁶ Second, cases from legal databases such as Westlaw are known to underreport jury decisions.⁴⁷ To some

related to trademark or service mark distinctiveness or likelihood of confusion were tracked from December 2011 to November 2012."); *see also* John R. Allison & Mark A. Lemley, *The (Unnoticed) Demise of the Doctrine of Equivalents*, 59 STAN. L. REV. 955, 963 (2007) ("[W]e collected every district court and court of appeals decision on the doctrine of equivalents that appeared in Westlaw . . ."). The number of cases this Article employs compares favorably with the norm, which ranges from less than 100 cases to 300 cases. *See* Hall & Wright, *supra* note 29, at 102 ("Of these 114 universal samples, only 11 coded more than 1000 cases, and 21 coded from 500 to 1000. Twenty-six of these projects coded fewer than 100 cases (with 13 of these fewer than 51), and 39 coded between 100 and 300.").

41. *See* Daryl Lim, Trademark Case Statistics (Dec. 17, 2021) (on file with the American University Law Review).

42. *Id.*

43. *Id.*

44. *Id.*

45. *See, e.g.,* Lim, *Doctrine of Equivalents*, *supra* note 35.

46. *Id.*

47. Allison & Lemley, *supra* note 40, at 963–64 ("The universe of all decisions is of course different from the universe of those reported in Westlaw, and in particular our study underreports jury decisions. But our focus on written decisions (both reported

degree, comparing it to other studies that employ similar methods to control for that feature can mitigate the effects of the underreporting.

Courts have found that “simple counts and percents are sufficient to document” a claim about case law trends, challenge conventional wisdom, or suggest further study issues.⁴⁸ The case-counting method codes the entire population of relevant cases.⁴⁹ Statistics are unnecessary to prove that sample cases are representative of a larger population.⁵⁰ Because the outcome of each case—the dependent variable of interest—has five possible categories, the most appropriate regression model would be a multinomial regression.⁵¹ However, each of the seven relevant factors also has five possible categories. Hence, the sample size required for a multivariate regression that would simultaneously test each category’s effect of each relevant factor is much larger than that of the current dataset.⁵² This Article refrained from presenting regression results and instead presented descriptive statistics. Specifically, it describes the observed distribution of case outcomes when conditioned on one or two relevant factors. Although this does not attest to a specific level of statistical significance, these values are still instructive for the reader.

Third, analysis of judicial opinions has well-known limitations.⁵³ Statistics fail to account for extralegal factors influencing outcomes, such as the state of the case record on appeal and judicial deliberations discussed in the opinion.⁵⁴ In addition, litigants may consider the expertise and reputation of the district court judge in deciding whether to appeal, introducing selection bias effects into the appellate

and unreported) allows us to parse the grounds for decision and the reasoning of the opinions.”).

48. Hall & Wright, *supra* note 29, at 118.

49. *Id.*

50. *Id.*

51. ALAN AGRESTI, CATEGORICAL DATA ANALYSIS 293 (3d ed. 2013).

52. *Id.*

53. See R. Polk Wagner & Lee Petherbridge, *Is the Federal Circuit Succeeding? An Empirical Assessment of Judicial Performance*, 152 U. PA. L. REV. 1105, 1128–29 (2004) (discussing unobserved reasoning, strategic behavior, and selection bias).

54. Harry T. Edwards & Michael A. Livermore, *Pitfalls of Empirical Studies that Attempt to Understand the Factors Affecting Appellate Decisionmaking*, 58 DUKE L.J. 1895, 1899 (2009).

data.⁵⁵ Moreover, most cases settle, so decided cases are a nonrandom subset of all cases.⁵⁶

Fourth, case content analysis trades depth for breadth. The complexity of trademark litigation also makes it difficult to generalize even from a study covering hundreds of cases.⁵⁷ Numbers do not reflect judicial rhetoric or more subtle clues about a judicial opinion's precedential value.⁵⁸ Case coding documents what judges do rather than draw normative implications from the observations.⁵⁹

Fifth, parties are not randomly distributed throughout the judicial districts.⁶⁰ Some district courts may hear more cases that eventually settle. Other courts may hear more cases where the parties file based on domicile. District court judges are therefore not assigned a random sample of patent lawsuits since they are assigned cases from the judicial district where they sit.⁶¹ Circumstances such as a particular judge or jury may cause a case to settle where the same case before another judge or jury could proceed to an appeal.⁶² This Article focuses on how appellate and lower courts interpret precedent. Those interpretations are not uniform and can never be so.⁶³

55. David Schwartz, *Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases*, 107 MICH. L. REV. 223, 243 (2008).

56. Jay P. Kesan & Gwendolyn G. Ball, *How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes*, 84 WASH. U. L. REV. 237, 273–74 (2006) (finding that between 65% and 68% of all patent cases filed in three particular years were resolved via settlement or a probable settlement).

57. See David L. Schwartz, *Explaining the Demise of the Doctrine of Equivalents*, 26 BERKELEY TECH. L.J. 1157, 1188 (2011) (“Because patent litigation as a whole is so complex, it is incredibly complicated to develop and test empirical models.”); Petherbridge, *Decline of the Doctrine of Equivalents*, *supra* note 36, at 1380 (noting biases inherent in this approach such as “unobserved reasoning, selection bias, and strategic behavior”).

58. Hall & Wright, *supra* note 29, at 87.

59. *Id.* at 97 (“Still, imperfect data must suffice because observing actual behaviors and gauging true attitudes would be impossible or cost-prohibitive. Similarly, even though judge-reported facts may not ‘purport to be the real facts,’ they are ‘near enough so that the savings in labor justifies the approximation.’”).

60. Kimberly A. Moore, *Forum Shopping in Patent Cases: Does Geographic Choice Affect Innovation?*, 79 N.C. L. REV. 889, 924–25 (2001).

61. Schwartz, *supra* note 55, at 242.

62. See *id.* at 242 n.119 (“[I]f [judges] have a really tough case, they can put tremendous pressure on the parties to settle so there won’t be an appealable order.” (alteration in original) (quoting Judge Richard A. Posner)).

63. See Jeffrey J. Rachlinski, *Does Empirical Evidence on the Civil Justice System Produce or Resolve Conflict?*, 65 DEPAUL L. REV. 635 (2016) (“[E]ven when the empirical scholars completely agree on the underlying facts, interpretation of the results can dramatically

Sixth, case outcomes are impacted by parties' factor-based calculation of a successful outcome. The Priest-Klein "selection hypothesis" predicts that, given various conditions, plaintiff win rates at trials should approach fifty percent because only the close cases survive settlement—or summary adjudication.⁶⁴ The hypothesis assumes parties have equal stakes in the litigation.⁶⁵ More recent studies cast the fifty-percent hypothesis in doubt, including those dealing specifically with intellectual property law.⁶⁶ Notably, "win rate" means the percentage of time one party (plaintiff or defendant) wins when a factor is decided in that party's favor, not the percentage of time that party wins when the factor is relevant. With these caveats in mind, this discussion turns to the theory underlying likelihood of confusion and the points of departure from conventional wisdom in practice.

B. Blends, Triggers, and Polaroid Factors

This Section opens by discussing the impact of blending technical trademarks and trade names in modern trademark law. It proceeds to introduce the likelihood of confusion factors before presenting a doctrinal and empirical analysis of intent, surveys, mark strength, and consumer sophistication, arguing that each, in turn, detracts from an accurate likelihood of confusion analysis. Finally, the Section closes by

differ. Empirical legal scholarship is still worth conducting, but the hope that it will resolve partisan debates in law is unrealistic.").

64. See George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 4–6, 17–18 (1984). Priest and Klein's fifty percent has been modified when there are different stakes involved. For example, if the plaintiff has more to win than the defendant has to lose. See Yoon-Ho Alex Lee & Daniel M. Klerman, *Updating Priest and Klein* 2 (Univ. S. Cal. Ctr. L. & Soc. Sci., Research Paper No. 15-21, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2619856).

65. Priest & Klein, *supra* note 64, at 24–29.

66. See e.g., Theodore Eisenberg, *Testing the Selection Effect: A New Theoretical Framework with Empirical Tests*, 19 J. LEGAL STUD. 337, 338–39 (1990) (testing the fifty-percent hypothesis and rejecting it as a description of all civil litigation); see also Mark A. Lemley & Colleen V. Chien, *Are the U.S. Patent Priority Rules Really Necessary?*, 54 HASTINGS L.J. 1299 (2003) (arguing that the Priest-Klein hypothesis is not borne out by the data in patent cases); Jason Rantanen, *Why Priest-Klein Cannot Apply to Individual Issues in Patent Cases*, (Univ. Iowa Legal Stud., Research Paper No. 12-15, 2012), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2132810 ("At best, the Priest-Klein hypothesis only applies to the selection of disputes, not the selection of individual issues. Due to the presence of multiple issues in patent cases, there is axiomatically no basis for inferring that a patentee would expect a 50 percent chance of winning on each one.").

explaining the impact of coherence-based reasoning on the likelihood of confusion factors.

1. *Blends and triggers*

Early trademark common law distinguished between trade names and technical trademarks.⁶⁷ Most trade name disputes involved rivals.⁶⁸ Unfair competition law governed these disputes and focused on directly competing uses diverting trade,⁶⁹ taking the form of passing off and reverse passing off business names.⁷⁰ In the twentieth century, courts blurred the distinction between technical trademarks and trade names, blending the most expansive aspects in favor of trademark owners.

In 1946, Congress “federalized” common law protection of trademarks used in interstate commerce with the Lanham Act.⁷¹ The Act codified this blended standard, requiring only that the unauthorized use be connected with goods or services.⁷² Trade names enjoyed the protection offered to technical trademarks as long as owners could show “secondary meaning.”⁷³ Cases interpreted this as an association by consumers with the source of the product that imbued trade names with an acquired distinctiveness.⁷⁴ The Act subsequently welded the two concepts, allowing all signs to acquire distinctiveness through secondary meaning.⁷⁵

67. See FRANK I. SCHECHTER, *THE HISTORICAL FOUNDATIONS OF THE LAW RELATING TO TRADE-MARKS* 161 (1925).

68. Edward S. Rogers, *The Lanham Act and the Social Function of Trademarks*, 14 *LAW & CONTEMP. PROB.* 173, 178–80 (1949).

69. See Mark P. McKenna, *The Normative Foundations of Trademark Law*, 82 *NOTRE DAME L. REV.* 1839, 1904 (2007) (noting “that courts only developed the likelihood of confusion factors after jettisoning the requirement of direct competition”).

70. See CORPORATE COUNSEL’S GUIDE TO UNFAIR COMPETITION § 3:31 (2021). “Passing off” occurs when defendants sell its goods with the plaintiff’s mark, whereas in “reverse passing off,” defendants sell plaintiff’s goods with the defendant’s trademark. *Id.*

71. *CAE, Inc. v. Clean Air Eng’g, Inc.*, 267 F.3d 660, 672 (7th Cir. 2001).

72. See Lanham Act, Pub. L. No. 79-489, § 32(1)(a), 60 Stat. 427, 437 (1946) (specifying that only goods and services fell under the purview of the Act).

73. See, e.g., Milton Handler & Charles Pickett, *Trade-Marks and Trade Names—An Analysis and Synthesis*, 30 *COLUM. L. REV.* 168, 168–69 (1930).

74. See *E. Columbia, Inc. v. Waldman*, 181 P.2d 865, 867 (Cal. 1947) (en banc).

75. Compare 15 U.S.C. § 1052(f) (2006) (allowing registration of a mark “which has become distinctive of the applicant’s goods in commerce”), with Trade-Mark Act of 1905, Pub. L. No. 58-84, § 5(b), 33 Stat. 724, 725–26 (providing that no mark that is distinguishable “from other goods of the same class shall be refused” trademark registration because of the nature of the mark).

As sellers expanded into adjacent product markets in the post-World War Two era, courts expanded the scope of protection to include complementary products and services.⁷⁶ Trademark scope could protect virtually anything that functions as an identifier of source—shapes, colors, smells, and sounds.⁷⁷ Congress included new types of protectable subject matter from technical trademarks to “anything . . . capable of carrying [source] meaning” as a potential trademark.⁷⁸ As a result, the likelihood of confusion standard became more complex. Although courts previously compared the marks, cases now require courts to consider a much broader range of information, including advertising slogans, product packaging, and product designs.

Figure 1: Trade Names, Technical Trademarks, and Modern Trademarks

	Trade Names	Technical Trademarks	Modern Trademarks
Distinctiveness	Sufficient if descriptive	Requires distinctiveness	Sufficient if descriptive
Intent	Intent required	Strict Liability	Intent optional
Harm	Actual harm required	Likelihood of harm sufficient	Likelihood of harm sufficient
Comparison	No	Yes	Optional
Injunction	Narrow	Broad	Broad

Contemporary empirical evidence from this Article’s dataset indicates that defendants win on the merits 26% of the time when the parties are rivals. Defendants win on the merits only 46% when they are not rivals, with rivals winning twice as often. These numbers show the impact of unfair competition in shaping modern trademark doctrine.

Congress subsequently amended the Act to remove the restriction on source confusion, allowing courts to consider other forms of

76. See PAMELA WALKER LAIRD, *ADVERTISING PROGRESS: AMERICAN BUSINESS AND THE RISE OF CONSUMER MARKETING* 31 (1998) (discussing post-war expansion of consumer products).

77. See Bone, *Notice Failure and Defenses*, *supra* note 11, at 1268.

78. *Qualitex Co. v. Jacobson Prods. Co.*, 514 U.S. 159, 162 (1995).

confusion in the infringement analysis.⁷⁹ Courts dutifully expanded the scope of confusion from purchasers to include non-purchasers (“post-sale confusion”) and allowed businesses to prohibit confusion over sponsorship or endorsement of goods and services.⁸⁰

Law and economics scholarship prompted this expansion, driven by a belief that stronger protection maximized wealth and, in turn, promoted economic efficiency.⁸¹ The resulting fusion infused unfair competition into trademark law and invited courts to find defendants’ marks infringing well before consumers purchased a product or service with the allegedly infringing mark, based on the idea that defendants misappropriated the plaintiff’s goodwill to appeal to consumers.⁸²

With new triggers, confusion can manifest itself in various ways. For instance, “forward confusion occurs when ‘the junior user attempts to trade on the senior[] user’s goodwill and reputation,’” misleading consumers to believe that the junior and senior user’s goods or services are related.⁸³ Similarly, reverse confusion occurs where consumers believe the junior user is the source of the senior user’s goods.⁸⁴ Whereas protection previously stopped at the shores of adjacent products, trademark law now allows even a pancake chain restaurant to attempt to prohibit an evangelical Christian organization from using

79. See S. REP. NO. 87-2107, at 4 (1962), *reprinted in* 1962 U.S.C.C.A.N. 2844, 2847.

80. Act of Oct. 9, 1962, Pub. L. No. 87-772, § 2, 76 Stat. 769, 769 (deleting the requirement that confusion be of purchasers as to the source of origin of such goods or services).

81. See *e.g.*, *W. T. Rogers Co. v. Keene*, 778 F.2d 334, 339 (7th Cir. 1985) (explaining that “competition is not impaired by giving each manufacturer a perpetual ‘monopoly’ of his identifying mark” if he has chosen a “distinctive” trademark where the available names are “for all practical purposes infinite”); see Landes & Posner, *supra* note 12, at 270–79 (advancing Chicago School economic theory within trademark law scope).

82. *Gibson Guitar Corp. v. Paul Reed Smith Guitars*, 423 F.3d 539, 549 (6th Cir. 2005).

83. *Fuel Clothing Co. v. Nike, Inc.*, 7 F. Supp. 3d 594, 610 (D.S.C. 2014) (quoting *Coryn Grp. II, LLC v. O.C. Seacrets, Inc.*, CIV WDQ-08-2764, 2010 WL 1375301, at *4 (D. Md. Mar. 30, 2010)).

84. *Kelly-Brown v. Winfrey*, 717 F.3d 295, 304–05 (2d Cir. 2013); *J.T. Colby & Co. v. Apple Inc.*, 586 F. App’x 8, 9 (2d Cir. 2014) (“The Lanham Act guards against this ‘reverse confusion’ to prevent ‘a larger, more powerful company [from] usurping the business identity of a smaller senior [trademark] user.’”).

a similar mark.⁸⁵ This development caused a jurisprudential disjuncture to occur.

While the statute had changed, earlier courts did not update the likelihood of confusion test, which had been designed to capture more than just source confusion.⁸⁶ As will be shown below, factors like consumer sophistication, the likelihood of expansion, and the marketing channels are of little assistance in evaluating whether a company's claim that it is the exclusive soda for sporting events in the minds of the consumers is true.⁸⁷ Worse, the multiple targets that the likelihood of confusion standard now addresses makes applying it even more unwieldy and unpredictable.⁸⁸

2. *The Polaroid factors*

The multifactor test for the likelihood of confusion attempts to provide analytical rigor to the complicated question of how consumers perceive different marks. Barton Beebe's 2006 empirical study revealed courts most frequently deployed the Second Circuit's test in *Polaroid Corp. v. Polaroid Electronics Corp.*⁸⁹ In that case, Judge Friendly articulated what became known as the eight *Polaroid* factors:

- (1) strength of the plaintiff's mark;
- (2) similarity of plaintiff's and defendant's marks;
- (3) competitive proximity of the products;
- (4) likelihood that plaintiff will "bridge the gap" and offer a product like a defendant's;
- (5) actual confusion between products;
- (6) good faith on the defendant's part;

85. First-Amended Complaint for Injunctive Relief for Federal Trademark Infringement and Dilution at 6, *IHOP IP, LLC v. Int'l House of Prayer*, No. CV10-6622 2010 WL 3775268 (C.D. Cal. Sept. 9, 2010).

86. See, e.g., *King of the Mountain Sports, Inc. v. Chrysler Corp.*, 185 F.3d 1084, 1090 (10th Cir. 1999) (discussing the three-part test to determine similarity between marks).

87. *Supreme Assembly, Ord. of Rainbow for Girls v. J.H. Ray Jewelry Co.*, 676 F.2d 1079, 1082 (5th Cir. 1982) ("other association"); *Caterpillar Inc. v. Walt Disney Co.*, 287 F. Supp. 2d 913, 918 (C.D. Ill. 2003) ("otherwise affiliated").

88. James Gibson, *Risk Aversion and Rights Accretion in Intellectual Property Law*, 116 *YALE L.J.* 882, 908 (2007) ("The case law on sponsorship and approval, however, is so ambiguous as to make it almost impossible to know ex ante whether a given use will be infringing.").

89. Beebe, *supra* note 6, at 1593 ("This is especially true in the Second Circuit where the multifactor test is most often applied and where appellate panels have repeatedly emphasized that the multifactor analysis must be exhaustive and explicit."); *Polaroid Corp. v. Polarad Elecs. Corp.*, 287 F.2d 492, 495 (2d Cir. 1961).

(7) quality of defendant's product; and

(8) sophistication of the buyers.⁹⁰

Confusion is more likely when an accused product contains multiple indicia of similarity.⁹¹ For instance, house brands typically include house marks, product-specific marks, product packaging, and color or configuration.⁹² Conversely, consumers are less likely to be confused when defendants copy only a few elements.⁹³ However, no single factor in the likelihood of confusion inquiry is determinative. Conventional wisdom teaches that courts need to undertake "a highly fact-intensive inquiry both as to the assessment of the evidence concerning each factor and as to the overall synthesis of factors and the evidence."⁹⁴

Trademark litigation is inherently impressionistic, particularly because actual confusion is rare. Sometimes, each side claims a numerically equal number of factors in their favor, leaving courts to assign weights.⁹⁵ Courts caught up in the swirl sloppily pepper their judgments with different operative terms to describe the same thing, including affiliation,⁹⁶ endorsement,⁹⁷ connection,⁹⁸ and whether the use produced confusion "of any kind."⁹⁹ As the Fifth Circuit bluntly put it, "Congress adopted an open-ended concept of confusion. . . . Any kind of confusion will now support an action for trademark infringement."¹⁰⁰

Unfortunately, courts in subsequent cases as well as businesses and their legal advisors struggle to determine the appropriate strength of

90. *Polaroid*, 287 F.2d at 495.

91. See Beebe, *supra* note 6, at 1625 (noting that judges "emphasize similarities over differences," but finding that "the degree of similarity of [] marks does not appear to significantly affect the outcome of the test").

92. See, e.g., *Bristol-Myers Squibb Co. v. McNeil-P.P.C., Inc.*, 973 F.2d 1033, 1042 (2d Cir. 1992).

93. George Miaoulis & Nancy D'Amato, *Consumer Confusion & Trademark Infringement*, 42 J. MKTG. 48, 54 (1978) (finding, in the context of competing goods, that the "primary cue for [] association [between two brands] was not the name but the visual appearance").

94. *Select Comfort Corp. v. Baxter*, 996 F.3d 925, 933–34 (8th Cir. 2021) ("We have repeatedly emphasized that no one factor is controlling and different factors will carry more weight in different settings.").

95. *Equitable Nat'l Life Ins. Co. v. AXA Equitable Life Ins. Co.*, 434 F. Supp. 3d 1227, 1252 (D. Utah 2020) ("Ultimately, while each side can claim three factors weigh in its favor, they do not weigh equally.").

96. *Pebble Beach Co. v. Tour 18 I Ltd.*, 155 F.3d 526, 544 (5th Cir. 1998).

97. *Id.*

98. *Id.* at 543.

99. *Syntex Lab'ys, Inc. v. Norwich Pharmacal Co.*, 437 F.2d 566, 568 (2d Cir. 1971).

100. *Armstrong Cork Co. v. World Carpets, Inc.*, 597 F.2d 496, 500 n.5 (5th Cir. 1979).

each factor, either alone or relative to other factors.¹⁰¹ Judges themselves admit the distinctions they make are often done on an “intuitive basis” rather than through “logical analysis.”¹⁰² Reporting on his dataset of cases, Beebe observed that “scattered among the circuits are factors that are clearly obsolete, redundant, or irrelevant, or, in the hands of an experienced judge or litigator, notoriously pliable.”¹⁰³

Like an untended garden, the likelihood of confusion standard has grown wild, with different circuit courts spinning off anywhere between six and thirteen factors.¹⁰⁴ Some circuits favor factors others ignore, and courts have called nearly every factor or factor combination the most important.¹⁰⁵ The reason for this may be divergent conceptions of trademark policy, with some courts focusing on unfair competition while others are concentrating on consumer confusion.¹⁰⁶

This Article reveals for the first time in the trademark literature that the Second Circuit’s *Polaroid* factors no longer dominate modern trademark jurisprudence. Instead, the Ninth Circuit’s *Sleekcraft* factors have edged out the *Polaroid* factors as those most frequently applied as the Ninth Circuit now has the most trademark infringement cases.¹⁰⁷ As a result, the Second Circuit is now the second most dominant circuit.

101. Joseph P. Liu, *Two-Factor Fair Use?*, 31 COLUM. J.L. & ARTS 571, 579 (2008) (“Under a multi-factor balancing test, it is difficult to register the relative strength of the factors.”); Eric Goldman, *Online Word of Mouth and its Implications for Trademark Law*, in TRADEMARK LAW AND THEORY: A HANDBOOK OF CONTEMPORARY RESEARCH 404, 415–16, 424 (Graeme B. Dinwoodie & Mark D. Janis eds., 2008) (“Assessing consumer confusion about product source is an inherently inexact process.”).

102. *Union Carbide Corp. v. Ever-Ready Inc.*, 531 F.2d 366, 379 (7th Cir. 1976).

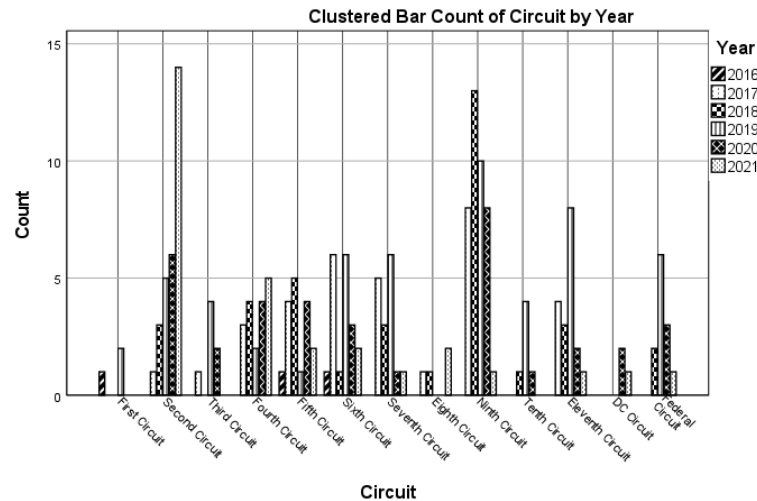
103. Beebe, *supra* note 6, at 1583–84.

104. *See infra*, Section I.B.

105. Beebe, *supra* note 6, at 1583.

106. Alejandro Mejías, *The Multifactor Test for Trademark Infringement from a European Perspective: A Path to Reform*, 54 IDEA 285, 314 (2014) (“[T]here is also divergence on how the factors are treated and employed.”); *see* Beebe, *supra* note 6, at 1596–97 (summarizing in chart form the different factors each circuit considers and reporting “substantial intercircuit variation in plaintiff multifactor test win rates.”).

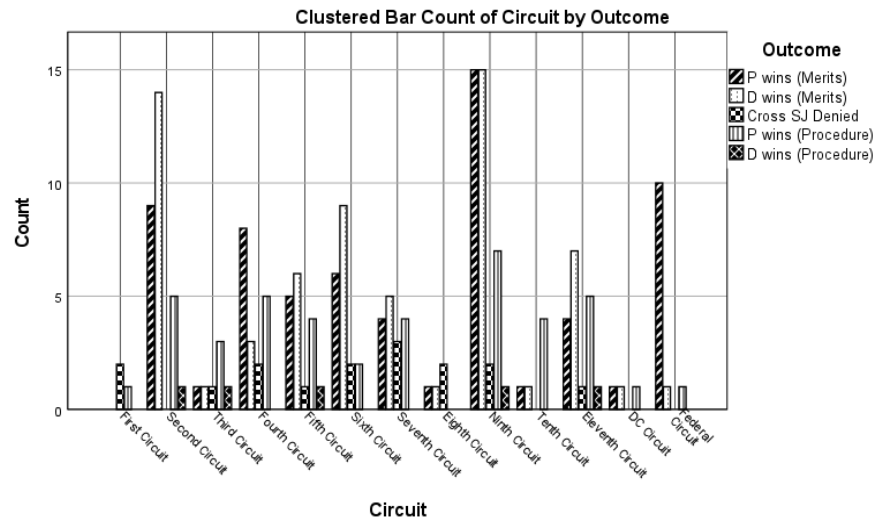
107. *AMF Inc. v. Sleekcraft Boats*, 599 F.2d 341, 350–51 (9th Cir. 1979).

Figure 2: Circuit Variances over Time

This shift may be significant for litigation strategy, especially because the Second Circuit was the most defendant-friendly circuit where plaintiffs' win rate was 31% and the defendant's win rate was 48%. By comparison, both plaintiff and defendant win rates were 38% at the Ninth Circuit. In contrast to the Second and Ninth Circuits, the Federal Circuit was the most plaintiff-friendly, with an 83% plaintiff win rate. Defendants there fare comparatively poorly, winning a mere 8% of cases.¹⁰⁸ The figure below shows the distribution of cases and outcomes across circuits.

108. It is possible that defendants fare so poorly at the Federal Circuit because they are likely on appeal from the Trademark Trial and Appeal Board. However, the data is inconclusive on this point and invites further study.

Figure 3: Circuit by Outcome



As a doctrinal matter, the difference may be less material. As it turns out, the Ninth Circuit's factors mirror those of the Second Circuit. The only difference between the two sets of factors is in linguistics as the Ninth Circuit considers marketing channels used to promote the products which is the same as the Second Circuit's consideration of the competitive proximity of the products and services.¹⁰⁹ Similarly, this Article shows that the Second Circuit's "quality of defendant's product" factor can be subsumed into the competitive proximity factor.¹¹⁰ As an empirical matter, both factors rarely appear in case reports, with the Ninth Circuit's "marketing channels" factor discussed in 13% of cases and the "quality" factor appearing in 11% of cases.¹¹¹

109. See e.g., *Las Vegas Sands Corp. v. Fan Yu Ming*, 360 F. Supp. 3d 1072, 1077 (D. Nev. 2019) (folding the two factors together).

110. See *infra* Section II.C.

111. Compare, for example, with mark similarity, which appeared in 85% of cases, competitive proximity in 73% of cases, and actual confusion in 74% of cases. See *infra* Appendix.

Figure 4: Comparing the Polaroid and Sleekcraft Factors

<i>Polaroid</i> Factors	<i>Sleekcraft</i> Factors
Strength of the plaintiff's mark	Strength of the plaintiff's mark
Similarity of plaintiff's and defendant's marks	Similarity of plaintiff's and defendant's marks
Competitive proximity of products or services	Competitive proximity of products or services
-	Marketing channels used to promote the products
Likelihood that plaintiff will "bridge the gap" and offer a product like a defendant's	Likelihood that either party will expand into new markets.
Actual confusion	Actual confusion
Defendant's good faith	Defendant's intent
Quality of defendant's product	-
Buyer sophistication	Type of goods and the carefulness of likely consumers

Substitution bias within each circuit's set of factors is particularly virulent when open-ended wording gives courts cover, as the Act does here.¹¹² Courts applying the Act took that opportunity and leaned into the likelihood of confusion factors like defendants' intent, survey evidence, and trademark strength, which were malleable and easy to wield to reach their desired outcomes.¹¹³ Savvy trademark attorneys also saw the opportunity to leverage more clever lawyering and focus less on the case's merits.¹¹⁴ Strikingly, Beebe's empirical study reported that intent and surveys were so heavily weighted that courts stampeded over other factors.¹¹⁵ Overreliance on these proxies results in a fundamentally flawed analysis. The next Section explains why.

112. See e.g., *SNA, Inc. v. Array*, 51 F. Supp. 2d 554, 562–63 (E.D. Pa. 1999), *aff'd sub nom. Silva v. Karlsen*, 259 F.3d 717 (3d Cir. 2001) (concluding that defendants' attempt to use metatags to "lure internet users to their site" was in bad faith).

113. See *infra* Section I.C.

114. See Beebe, *supra* note 6, at 1581 (suggesting this problem exists even with true source confusion cases because outcomes tend to be driven by the court's focus on intent).

115. *Id.* at 1607.

C. Missing the Point on Consumer Confusion

What constitutes “confusion” is highly subjective and difficult to evaluate. Proxies like intent, survey evidence, mark strength, and consumer sophistication fail to incorporate real-world purchasing conditions or are better considered within other factors. Trademark infringement is fundamentally flawed if the likelihood of confusion turns on these proxies.

1. Intent

Likelihood of confusion’s good faith or intent factor examines whether defendants sought to benefit from plaintiffs’ goodwill.¹¹⁶ All circuits but the Federal Circuit recognized this as a major factor in finding liability.¹¹⁷ “In analyzing whether a defendant has acted in bad faith, the question is whether the defendant attempted ‘to exploit the good will and reputation of a senior user by adopting the mark with the intent to sow confusion between the two companies’ products.’”¹¹⁸

Courts recognize that intentional copying may not indicate that the defendant attempted to capitalize on the plaintiff’s trademark or trade dress.¹¹⁹ However, there may be legitimate reasons to copy or imitate the primary features of another company’s product. These include functional features that have economic benefits without any secondary meaning.¹²⁰ In doing so, courts “want competitors to be inspired by—and to improve on—the findings of their predecessors.”¹²¹ Therefore, it is a “nefarious variety of passing off—the kind that confuses consumers and exploits a competitor’s established goodwill—that trademark law is prepared to prevent.”¹²²

Stating the distinction is easy in theory, hard in practice. Cases in the dataset reveal divergent views on when defendants cross the line. Some courts are prepared to exculpate defendants if they had no intent to confuse consumers.¹²³ Indeed, one court commended “upcycling,” or

116. *Sicilia Di R. Beibow & Co. v. Cox*, 732 F.2d 417, 431 (5th Cir. 1984) (discussing how the proper test focuses mainly on intent).

117. *See Beebe*, *supra* note 6, at 1589–90.

118. *Tiffany & Co. v. Costco Wholesale Corp.*, 971 F.3d 74, 88 (2d Cir. 2020).

119. *See Beebe*, *supra* note 6, at 1630.

120. *Fuddruckers, Inc. v. Doc’s B.R. Others, Inc.*, 826 F.2d 837, 844–45 (9th Cir. 1987).

121. *Vital Pharms., Inc. v. Monster Energy Co.*, No. 19-60809-CIV, 2021 WL 3371942, at *44 (S.D. Fla. Aug. 3, 2021).

122. *Id.*

123. *Id.* at *53 (“If a defendant intentionally copies an aspect of the plaintiff’s product, but not with intent to confuse consumers, then the defendant’s intent has

“restoring previously nonfunctional antique watch movements and parts,” as good faith.¹²⁴ The court used this reasoning despite the defendant’s intent to benefit from displaying the plaintiff’s mark, though intending to capitalize on its historical significance rather than its modern-day reputation.¹²⁵

Others courts stand ready to pin the defendant down on a lower negligence standard for failure to exercise due diligence.¹²⁶ Yet some will find against the defendant on an attempt standard, even without proof that actual confusion resulted from it,¹²⁷ and “some courts find evidence of bad faith even where they conclude the defendant did not choose its mark purposely to promote confusion.”¹²⁸ Unsurprisingly, these courts emphatically state that a defendant’s lack of intent is generally not relevant to the likelihood of consumer confusion.¹²⁹ Yet, surprisingly, the presence of intent may not be decisive either. For example, in one case, the court expressed that even when there is explicit evidence of bad faith, that factor alone should not determine the outcome of a case.¹³⁰ Instead, a defendant’s “[b]ad faith and intent to deceive are relevant to the extent that they add to the likelihood

little bearing on the ultimate question: whether the allegedly infringing product is likely to confuse consumers.”); *QuikTrip W., Inc. v. Weigel Stores, Inc.*, 984 F.3d 1031, 1036 (Fed. Cir. 2021) (“[T]he ‘only relevant intent is intent to confuse. There is a considerable difference between an intent to copy and an intent to deceive.’”).

124. *Hamilton Int’l Ltd. v. Vortic LLC*, 486 F. Supp. 3d 657, 667–68 (S.D.N.Y. 2020), *aff’d*, 13 F.4th 264, 268 (2d Cir. 2021) (“The Court credits this testimony, concluding that he did not intend to cause consumer confusion but rather sought to ‘preserve American history’ by salvaging and restoring the hearts of antique pocket watches.”).

125. *Id.* at 668; *see also* *Champion Spark Plug Co. v. Sanders*, 331 U.S. 125, 130 (1947) (noting that it can be “wholly permissible” that the “second-hand dealer gets some advantage from the trademark”).

126. *AWGI, LLC v. Atlas Trucking Co.*, 998 F.3d 258, 268 (6th Cir. 2021) (finding the intent factor irrelevant); *see also* *Ironhawk Techs., Inc. v. Dropbox, Inc.*, 994 F.3d 1107, 1124 (9th Cir. 2021), *amended by* 2 F.4th 1150 (9th Cir. 2021) (“This factor ‘favors the plaintiff where the alleged infringer adopted his mark with knowledge, actual or constructive, that it was another’s trademark.’”).

127. *ServPro Intell. Prop., Inc. v. Blanton*, 451 F. Supp. 3d 710, 727 (W.D. Ky. 2020) (“Courts have held that ‘[i]f a party chooses a mark with the intent of causing confusion, *that fact alone* may be sufficient to justify an inference of confusing similarity.’” (emphasis added) (quoting *Homeowners Grp., Inc. v. Home Mktg. Specialists, Inc.*, 931 F.2d 1100, 1111 (6th Cir. 1991))).

128. *Equitable Nat’l Life Ins. Co. v. AXA Equitable Life Ins. Co.*, 434 F. Supp. 3d 1227, 1248 (D. Utah 2020).

129. *ServPro Intell. Prop., Inc.*, 451 F. Supp. 3d, at 727; *Variety Stores, Inc. v. Wal-Mart Stores, Inc.*, 888 F.3d 651, 665 (4th Cir. 2021) (“Bad faith may be inferred from [Walmart’s] actual or constructive knowledge of [Variety’s] mark.”).

130. *Car-Freshner Corp. v. Am. Covers, LLC*, 980 F.3d 314, 333 (2d Cir. 2020).

that the accused infringer will achieve its objective of consumer confusion.”¹³¹

Beverly Pattishall suggested that factoring in intent makes outcomes more predictable.¹³² This inference makes it easier to determine the state of mind of one person, the defendant, than to forecast the perceptions of the consumer group.¹³³ Predictability is good, but the result may not be. Intent inherently focuses on the wrong goalpost. Merely because the defendant’s mental state is easier to discern than the perception of the consuming public does not make that factor more relevant to the inquiry. As Kelly Collins warned, “[t]his is dangerous because mere ‘copying’ is not always impermissible.”¹³⁴ The law encourages reusing generic or functional marks “as a part of our competitive economic system.”¹³⁵ For this reason, she argues that the relevant intent is the one to confuse and not merely to copy.¹³⁶

Another reason to abandon intent is that it muddies jurisprudential waters caused by further fusion of trade name and technical trademark jurisprudence. Courts typically require intent when dealing with non-inherently distinctive marks.¹³⁷ Courts have either presumed intent or dispensed with it for inherently distinctive marks.¹³⁸

Alejandro Mejías explained that intent is irrelevant because the focus “is not what the defendant intended to do, but whether his mark is likely to be confusingly similar for the relevant public.”¹³⁹ Very few courts acknowledge this much.¹⁴⁰ Judges may like intent because it makes their job easier, and the outcome feels more just. However,

131. *Id.*

132. See Beverly W. Pattishall, *The Impact of Intent in Trade Identity Cases*, 60 TRADEMARK REP. 575, 579–80 (1970).

133. *Id.* at 577.

134. Kelly Collins, Comment, *Intending to Confuse: Why Preponderance Is the Proper Burden of Proof for Intentional Trademark Infringements Under the Lanham Act*, 67 OKLA. L. REV. 73, 87 (2014).

135. *Id.*

136. *Id.* at 87–88 (“This would better serve the purposes of the Lanham Act and safeguard innocent conduct from triggering liability.”).

137. But see 4 MCCARTHY, *supra* note 2, § 23:106 (explaining that proof of intent is merely evidence relevant to whether confusion is likely).

138. See *Cashmere & Camel Hair Mfrs. Inst. v. Saks Fifth Ave.*, 284 F.3d 302, 317 (1st Cir. 2002) (explaining that where there is evidence of intentional deceit the presumption is clear); *Res. Devs., Inc. v. Statue of Liberty-Ellis Island Found., Inc.*, 926 F.2d 134, 140 (2d Cir. 1991).

139. Mejías, *supra* note 106, at 349.

140. See, e.g., *Virgin Enters. Ltd. v. Nawab*, 335 F.3d 141, 151 (2d Cir. 2003) (explaining that intent is not “of high relevance to the issue of likelihood of confusion” because “[i]t does not bear directly on whether consumers are likely to be confused”).

intent is irrelevant to technical trademark infringement. Recall from Section I.B that technical trademark infringement focuses on the consequences of the defendant's act and not on their intent.¹⁴¹ In contrast, trade name infringement focuses on defendants' desired outcomes, irrespective of consumer confusion.¹⁴²

Pinning the likelihood of confusion on free-riding becomes problematic because free-riding is ultimately a concept searching for meaning.¹⁴³ The Act does not require proof of intent. Trademark infringement is, after all, a strict liability offense.¹⁴⁴ As the Sixth Circuit opined, the better view is to consider intent only *after* other likelihood of confusion factors indicate liability.¹⁴⁵ Intent may go to aggravated remedies, but it should be irrelevant to the question of guilt. As Beebe put it, "if trademark law seeks to prevent commercial immorality, then it should do so explicitly. An injunction should issue, and damages be granted on that basis alone, and not on the basis of possibly distorted findings of fact as to the likelihood of consumer confusion."¹⁴⁶

Beebe found that despite the disconnect between the defendant's intent and consumer confusion, it stampedes the other factors.¹⁴⁷ The effect is powerful—a "nearly un-rebuttable presumption of a likelihood of confusion"¹⁴⁸ roughly 97% of the time,¹⁴⁹ making it "arguably the single most important confusion factor in use today."¹⁵⁰

141. Rogers, *supra* note 68, at 178 (explaining the origins of trademark law).

142. See, e.g., *Visible Sys. Corp. v. Unisys Corp.*, 551 F.3d 65, 73 (1st Cir. 2008) (explaining that sometimes there is a likelihood of confusion in industries regardless of intent).

143. See, e.g., *Ty Inc. v. Perryman*, 306 F.3d 509, 512 (7th Cir. 2002) (rejecting sponsorship dilution claim because "in that attenuated sense of free riding, almost everyone in business is free riding").

144. See *Taubman Co. v. Webfeats*, 319 F.3d 770, 775 (6th Cir. 2003) (recognizing that the Lanham Act is a "strict liability statute"); see also Rebecca Tushnet, *Running the Gamut From A to B: Federal Trademark and Federal False Advertising Law*, 159 U. PA. L. REV. 1305, 1310 (2011) (noting that federal courts have interpreted trademark as a strict liability offense); Robert G. Bone, *Enforcement Costs and Trademark Puzzles*, 90 VA. L. REV. 2099, 2109 (2004) (referring to trademark infringement as a form of strict liability).

145. See, e.g., *Taubman Co.*, 319 F.3d at 775 ("[T]he proper inquiry is not one of intent. In that sense, the Lanham Act is a strict liability statute. If consumers are confused by an infringing mark, the offender's motives are largely irrelevant.").

146. Beebe, *supra* note 6, at 1631.

147. *Id.* at 1621.

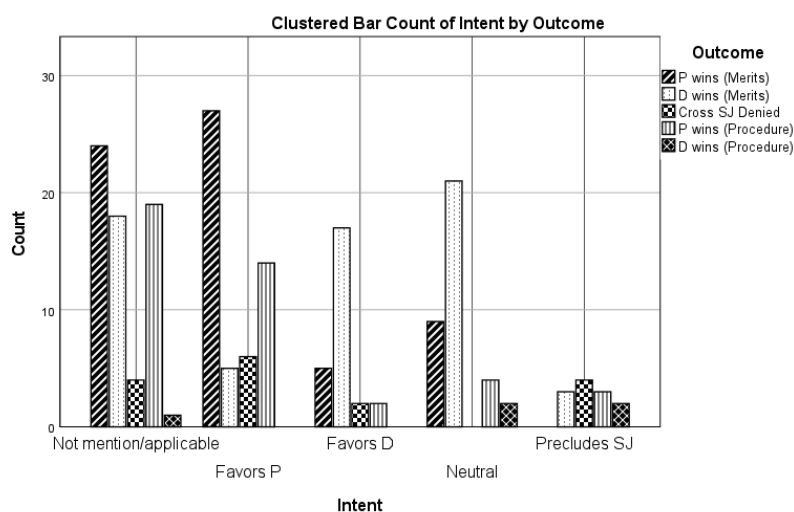
148. *Id.* at 1628.

149. *Id.* ("[The data] suggest that a finding of bad faith intent creates, if not in doctrine, then at least in practice, a nearly un-rebuttable presumption of a likelihood of confusion.").

150. Blake Tierney, *Missing the Mark: The Misplaced Reliance on Intent in Modern Trademark Law*, 19 TEX. INTELL. PROP. L.J. 229, 236 (2011).

This Article reports that intent appeared in two-thirds of the cases studied and was deemed neutral 19% of the time. In 27% of all cases, courts favored plaintiffs on the intent factor. When they did, plaintiffs won 52% of the time. In 14% of all cases, the courts favored defendants on the intent factor. When the court favored defendants, they won 65% of the time.

Figure 5: Intent by Outcome



Qualitatively, cases in the dataset show that intent bears a minimal impact on results.¹⁵¹ The reason is that “an intent to confuse customers is not required for a finding of trademark infringement.”¹⁵² Nonetheless, the intensely fact-specific nature of intent can trip up parties seeking speedy resolution of the dispute. As one court in the dataset put it, “[i]ssues of bad or good faith ‘are generally ill-suited for disposition on summary judgment.’”¹⁵³ The practice is longstanding,

151. *Equitable Nat’l Life Ins. Co. v. AXA Equitable Life Ins. Co.*, 434 F. Supp. 3d 1227, 1248 (D. Utah 2020) (“Although this factor weighs in AXA’s favor, its impact is minimal.”); *GoTo.com, Inc. v. Walt Disney Co.*, 202 F.3d 1199, 1208 (9th Cir. 2000) (emphasizing “the minimal importance of the intent factor”).

152. *GoTo.com*, 202 F.3d at 1208 (citation omitted).

153. *RVC Floor Decor, Ltd. v. Floor & Decor Outlets of Am., Inc.*, 527 F. Supp. 3d 305, 327 (E.D.N.Y. 2021); *see also* *Zamfir v. Casperlabs, LLC*, 528 F. Supp. 3d 1136, 1145 (S.D. Cal. 2021) (“These unresolved factual questions complicate the issue of Defendant’s intent in choosing the mark.”).

with courts preferring to leave it to juries to settle the matter.¹⁵⁴ Surprisingly, at least one court insisted on a jury trial even when the marks in question were identical due to the inherently subjective nature of the inquiry.¹⁵⁵

The dataset shows that 6% of cases expressly precluded summary judgment based on the intent factor. That figure may seem low, but it is considerably higher than any of the other factors: mark similarity (4%), buyer sophistication (3%), actual confusion (2%), mark strength (1%), competitive proximity (1%), “bridging the gap” (0%), and quality (0%).

Eliminating intent allows a more focused inquiry into the likelihood of confusion rather than the commercial immorality of defendants. As a practical matter, it frees parties from costly discovery and allows the court to grant summary judgment more frequently.¹⁵⁶ Judges can also dispose of cases more easily without trial, and it is less likely that defendants will be subject to vexatious suits based on the nebulous aspersions of intent.¹⁵⁷

While the “ordinary consumer” is central to the infringement analysis, it remains poorly theorized.¹⁵⁸ In patent cases, courts benefit from expert testimony.¹⁵⁹ Perhaps this is because the subject matter of patent disputes is by nature technologically challenging, defaulting those involved to accept, even expect, expert assistance. But trademark courts must investigate confusion without evidence that any consumers

154. See *EMI Catalogue P’ship v. Hill, Holliday, Connors, Cosmopolos Inc.*, 228 F.3d 56, 67–68 (2d Cir. 2000) (“Because the issue goes to defendants’ intent, it ‘is best left in the hands of the trier of fact.’”).

155. *Tiffany & Co. v. Costco Wholesale Corp.*, 971 F.3d 74, 88 (2d Cir. 2020) (“And as we have consistently observed, ‘subjective issues such as good faith are singularly inappropriate for determination on summary judgment.’”).

156. 10B CHARLES ALAN WRIGHT ET AL., *FEDERAL PRACTICE AND PROCEDURE (CIVIL)* § 2730 n.3 (3d ed. 2015) (“Questions of intent, which involve intangible factors including witness credibility, are matters for the consideration of the fact finder after a full trial and are not for resolution by summary judgment.”).

157. Thomas L. Casagrande, *A Verdict for Your Thoughts? Why an Accused Trademark Infringer’s Intent Has No Place in Likelihood of Confusion Analysis*, 101 TRADEMARK REP. 1447, 1455 (2011) (proposing an elimination of intent as a factor to be considered in determining trademark infringement).

158. See e.g., Thomas R. Lee et al., *Trademarks, Consumer Psychology, and the Sophisticated Consumer*, 57 EMORY L.J. 575, 575 (2008) (“[N]either courts nor commentators have made any serious attempt to develop a framework for understanding the conditions that may affect the attention that can be expected to be given to a particular purchase.”).

159. Douglas G. Smith, *The Increasing Use of Challenges to Expert Evidence Under Daubert and Rule 702 in Patent Litigation*, 22 J. INTELL. PROP. L. 345, 354 (2015).

were confused, imagining consumers' likely experience as filtered through the parties' competing interests. This notional consumer is "neither savant nor dolt."¹⁶⁰ One who "lacks special competency with reference to the matter at hand but has and exercises a normal measure of the layman's common sense and judgment."¹⁶¹ Instead, courts rely on surveys, mark strength, and consumer sophistication to determine the likelihood of confusion. But, like intent, none of these factors provide a good proxy. The Sections below explain why.

2. *Surveys*

Surveys attempt to measure whether consumers believe that the plaintiff's mark is the source of the alleged infringer's product or whether it sponsors or approves it.¹⁶² Plaintiffs may provide survey evidence that an appreciable number of relevant consumers are likely to be confused.¹⁶³ According to a case in the dataset, survey evidence is not a prerequisite for establishing public recognition, but "it is the most persuasive evidence of it."¹⁶⁴

Surveys present respondents with defendants' marks and measure consumers' reactions in the context that consumers encounter the mark in question.¹⁶⁵ Proof of marketing supports broad public recognition.¹⁶⁶ They typically involve control groups to show causality between the defendants' mark and consumer confusion.¹⁶⁷

In theory, a survey needs to pass muster under the Federal Rules of Evidence, which requires considering the "validity of the techniques employed."¹⁶⁸ Courts can bar significantly flawed surveys as evidence

160. Kraft Foods Grp. Brands LLC v. Cracker Barrel Old Country Store, Inc., 735 F.3d 735, 743 (7th Cir. 2013).

161. United States v. 88 Cases, More or Less, Containing Bireley's Orange Beverage, 187 F.2d 967, 971 (3d Cir. 1951).

162. 3 ANNE GILSON LALONDE, GILSON ON TRADEMARKS § 8.03 (2021).

163. See 6 MCCARTHY, *supra* note 2, § 32:158.

164. Kibler v. Hall, 843 F.3d 1068, 1074 (6th Cir. 2016).

165. Shari Seidman Diamond & David J. Franklyn, *Trademark Surveys: An Undulating Path*, 92 TEX. L. REV. 2029, 2037 (2014).

166. Therma-Scan, Inc. v. Thermoscan, Inc., 295 F.3d 623, 632 (6th Cir. 2002).

167. See, e.g., Bracco Diagnostics, Inc. v. Amersham Health, Inc., 627 F. Supp. 2d 384, 448 (D.N.J. 2009) (criticizing a survey's design for failure to use "a control mechanism").

168. Shari Seidman Diamond, *Reference Guide on Survey Research*, in REFERENCE MANUAL ON SCI. EVIDENCE 359, 364 (3d ed. 2011).

when they are more prejudicial than probative¹⁶⁹ or deemed unreliable.¹⁷⁰

The problem is that commentators and courts alike acknowledge that surveys are often unreliable and expensive, costing hundreds of thousands of dollars.¹⁷¹ Courts routinely attack the representativeness of the survey from a parade of cherry-picked witnesses and extrapolate a standard of what consumers generally believe.¹⁷² The inexact science of assessing trademark strength causes judges to rely upon or reject surveys based on whether the results agree with their subjective impressions.¹⁷³ As a result, judicial unease with surveys sometimes bubbles to the surface, with Judge Richard Posner remarking that “no doubt there are other tricks of the survey researcher’s black arts that we have missed.”¹⁷⁴

Constructing a robust survey is dauntingly hard. Surveys need to employ a control¹⁷⁵ and calculate noise.¹⁷⁶ As an indication of the

169. *Citizens Fin. Group, Inc. v. Citizens Nat’l Bank*, 383 F.3d 110, 120 (3d Cir. 2004).

170. *Id.*

171. *Nabisco, Inc. v. PF Brands, Inc.*, 191 F.3d 208, 224 (2d Cir. 1999) (“[C]onsumer surveys . . . are expensive, time-consuming and not immune to manipulation.”); *Johnson v. Revenue Mgmt. Corp.*, 169 F.3d 1057, 1063 (7th Cir. 1999) (“[S]urvey evidence in trademark and trade dress cases can be very costly.”); see also Robert H. Thornburg, *Trademark Survey Evidence: Review of Current Trends in the Ninth Circuit*, 21 SANTA CLARA COMPUT. & HIGH TECH. L.J. 715, 717 (2005) (“[T]he most basic of surveys cost[s] in the hundreds of thousands of dollars.”).

172. *But cf. Citizens Fin. Grp.*, 383 F.3d at 122 (“In general, ‘actual confusion’ evidence collected by employees of a party in a trademark action must be viewed with skepticism because it tends to be biased or self-serving.”).

173. 6 MCCARTHY, *supra* note 2, § 32:196 (“Since an estimation of the probable mental reactions and associations of the buying public is not a science, there is always the temptation to decide on the basis of a ‘hunch.’ That is, the trier of fact (or any human being) would rather extrapolate from his or her own subjective impressions than extrapolate from some hard evidence of other persons’ subjective impressions—especially if the two do not agree.”); see also Peter Weiss, *The Use of Survey Evidence in Trademark Litigation: Science, Art or Confidence Game?*, 80 TRADEMARK REP. 71, 83 (1990) (“[A] reading of the many cases in which either great weight or little weight was given to survey evidence will, I feel reasonably certain, lead most objective analysts to the conclusion that, while some surveys went down because they were indeed ‘seriously flawed,’ many others either stayed up or went down depending on the result which the judges wanted to reach.”).

174. *Indianapolis Colts, Inc. v. Metro. Balt. Football Club Ltd. P’ship*, 34 F.3d 410, 416 (7th Cir. 1994).

175. 6 MCCARTHY, *supra* note 2, § 32:187 (“Courts have held that a survey that fails to use a control may be given less weight or even excluded from evidence altogether.”).

176. DANIEL KAHNEMAN, OLIVER SIBONY & CASS SUNSTEIN, *NOISE: A FLAW IN HUMAN JUDGMENT* 488 (2021) (“Noise is variability in judgments that should be identical.”).

treacherousness of this task, one court dismissed an expert witness who authored a book on the very subject of conducting trademark surveys for producing a “useless” survey.¹⁷⁷ As a matter of justice between the parties, the staggering costs of surveys put defendants at a disadvantage. Robert Bone explained that “[p]roving a high [likelihood of confusion] puts a premium on surveys and expert testimony and is likely to require extensive discovery, all of which will increase direct litigation costs and strengthen a trademark owner’s ability to leverage cease-and-desist threats in frivolous and weak cases.”¹⁷⁸

Qualitatively, cases in the dataset warn that surveys only represent circumstantial evidence of actual confusion, providing an experimental environment, not real consumers making mistaken purchases.¹⁷⁹ As one court put it, “[a]necdotal evidence can be more direct evidence of actual confusion and so is ‘both relevant and probative.’”¹⁸⁰ Another court “noted a trend away from according great weight to survey evidence,” and afforded the survey no weight.¹⁸¹ Unlike actual confusion, that court explained that “survey evidence is circumstantial, not direct, evidence of the likelihood of confusion. Surveys do not measure the degree of actual confusion by real consumers making mistaken purchases.”¹⁸² Accordingly, the court faulted the survey for “depart[ing] from real-market conditions in a way that was both biased and misleading.”¹⁸³

177. *Vital Pharms., Inc. v. Monster Energy Co.*, No. 19-60809-CIV, 2021 WL 3371942, at *65 (S.D. Fla. Aug. 3, 2021).

178. See Bone, *Notice Failure and Defenses*, *supra* note 11, at 1269 n.110.

179. *Vital Pharms., Inc.*, 2021 WL 3371942, at *55 (“[S]urvey evidence is not direct evidence of customer confusion in the real marketplace.”).

180. *Juul Labs, Inc. v. 4X PODS*, 509 F. Supp. 3d 52, 66 (D.N.J. 2020).

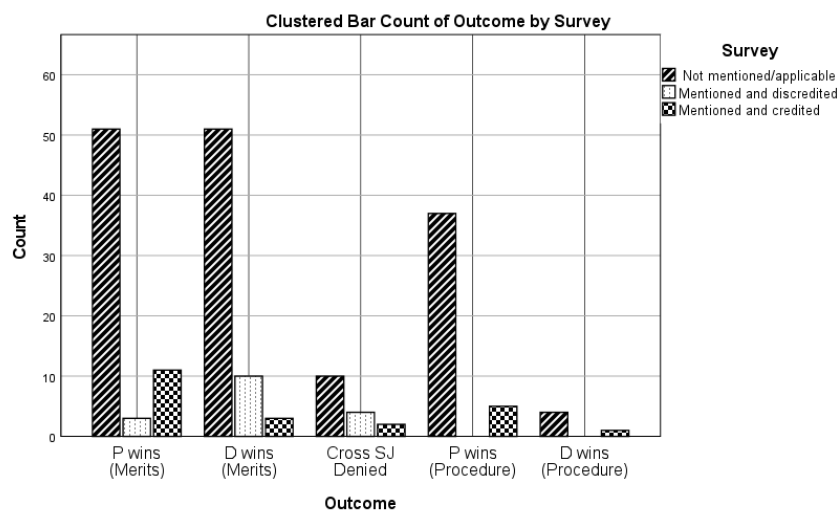
181. *Vital Pharms., Inc.*, 2021 WL 3371942, at *55.

182. *Id.* at *64.

183. *Id.* at *65 (“[E]vidence at trial confirmed the obvious: that the artificial coolers Mr. Berger showed his survey participants looked nothing like the coolers consumers would encounter in real stores.”); see, e.g., *Citizens Banking Corp. v. Citizens Fin. Grp., Inc.*, 320 F. App’x 341, 348 n.4 (6th Cir. 2009) (minimizing the weight of a confusion survey because it “failed to mimic the purchase conditions”); *Coherent, Inc. v. Coherent Techs., Inc.*, 935 F.2d 1122, 1126 (10th Cir. 1991) (affirming the district court’s finding that the “survey did not show actual confusion because it failed to simulate decisions in the marketplace”); *Am. Footwear Corp. v. Gen. Footwear Co.*, 609 F.2d 655, 661 (2d Cir. 1979) (finding that “the critical defect in this survey was the failure to conduct it under actual marketing conditions”—and so the “district court’s rejection of this survey evidence was not clearly erroneous”); 4 MCCARTHY, *supra* note 2, § 23:2.50 (stating that a survey is only evidence of confusion if “the survey mirrors the real world setting which can create an instance of actual confusion”).

Scholars also warn against placing a premium on surveys. According to Beebe, “the conventional view of the utility of survey evidence may be incorrect”: only 20% of the cases he reviewed addressed survey evidence, 10% credited survey evidence, and 7% ruled in favor of the outcome that the credited survey evidence favored.¹⁸⁴ This dataset shows a near-identical result fifteen years later. Of the 20% of cases that addressed survey evidence, 12% credited survey evidence, and 6% ruled in favor of the outcome that the credited survey evidence favored.

Figure 6: Outcome by Survey Evidence



As with intent, there is a certain circular irony to the whole exercise regarding surveys. Courts rely on surveys only to support conclusions that judges reach using other factors. The analysis also works backward—faced with survey evidence showing a likelihood of confusion, judges may regard the marks as more similar than they might have appeared in the absence of the survey.¹⁸⁵ As Peter Weiss remarked, “[o]ne might sum it all up by saying that the function of surveys in trademark litigation is to plumb the minds of the public to make up the minds of the judges.”¹⁸⁶ Dispensing with surveys and

184. Beebe, *supra* note 6, at 1641. A more recent study that expanded Beebe’s dataset found that only about 17% of cases addressed survey evidence. See Robert C. Bird & Joel H. Steckel, *The Role of Consumer Surveys in Trademark Infringement: Empirical Evidence from the Federal Courts*, 14 U. PA. J. BUS. L. 1013, 1035 (2012).

185. Diamond & Franklyn, *supra* note 165, at 2043.

186. Weiss, *supra* note 173, at 86.

relying on the court's judgment would not only be cheaper and simpler, but it would also be the intellectually honest thing to do.

Surveys sometimes overlap with trademark strength since parties may use the former to measure the potency of a mark's goodwill and its worthiness of protection.¹⁸⁷ Known as the *Abercrombie* spectrum, generic and descriptive marks are not distinctive, suggestive marks are marginally distinctive, while arbitrary or fanciful marks are inherently distinctive.¹⁸⁸ Trademark strength is usually the first factor courts consider.¹⁸⁹

3. *Mark strength*

A mark's distinctiveness is its uniqueness in denoting a product. Marks may be fanciful, arbitrary, suggestive, descriptive, or generic from most to least distinctive.¹⁹⁰ Generic terms are unprotectable and descriptive ones are protectable only when buyers view them as distinctive of a unique source.¹⁹¹ Evaluating the strength of a mark requires the fact finder to evaluate several factors: its degree of inherent distinctiveness, its "conceptual strength," its distinctiveness in the marketplace, and its "commercial strength."¹⁹² Unlike conceptual strength, commercial strength considers advertising expenditures, consumer studies linking the mark to a source, sales success, unsolicited media coverage of the product, attempts to plagiarize the mark, and the length and exclusivity of the mark's use.¹⁹³

The dataset reveals that mark strength comes up in 70% of the cases. In 47% of all cases, the courts favored plaintiffs on the mark strength factor. When they did, plaintiffs won 46% of the time. In 15% of all

187. Beebe, *supra* note 6, at 1646 ("In trademark law, the question is always of consumer perception in the marketplace rather than judicial perception in the courtroom.").

188. *Abercrombie & Fitch Co. v. Hunting World, Inc.*, 537 F.2d 4, 9 (2d Cir. 1976).

189. *See, e.g., Welding Servs., Inc. v. Forman*, 509 F.3d 1351, 1361 (11th Cir. 2007) ("The stronger or more distinctive a trademark or service mark, the greater the likelihood of confusion . . ."); Barton Beebe & C. Scott Hemphill, *The Scope of Strong Marks: Should Trademark Law Protect the Strong More Than the Weak?*, 92 N.Y.U. L. REV. 1339, 1349 n.40 (2017) ("Strength is the first factor in the Second, Fourth, Fifth, Sixth, Eighth, Ninth, and Eleventh Circuits, the second factor in the Third Circuit, and the last factor in the First and Tenth Circuits."). Courts consider design marks under the *Seabrook* factors. *See Seabrook Foods, Inc. v. Bar-Well Foods, Ltd.*, 568 F.2d 1342 (C.C.P.A. 1977).

190. *U.S. Pat. & Trademark Off. v. Booking.com B. V.*, 140 S. Ct. 2298, 2302 (2020).

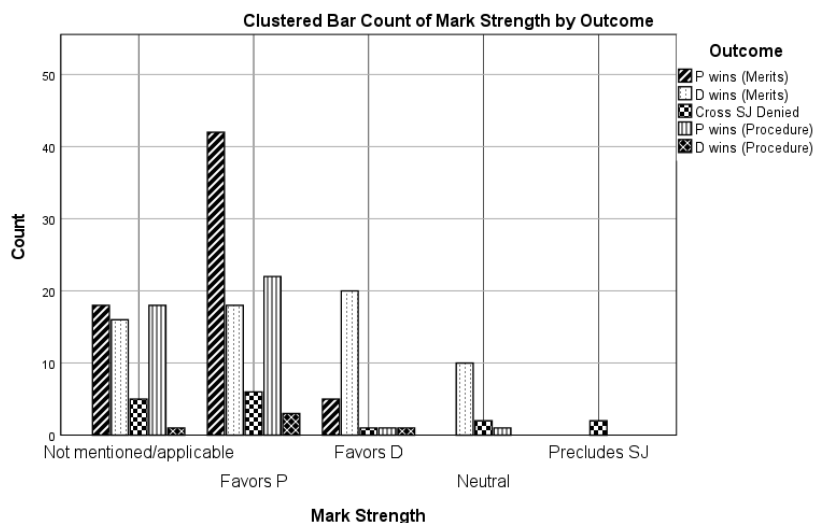
191. *Id.* at 2303.

192. Ouellette, *supra* note 40, at 353.

193. *Variety Stores, Inc. v. Walmart Inc.*, 852 F. App'x 711, 719 (4th Cir. 2021).

cases, the courts favored defendants on the mark strength factor. When the court favored defendants, they won 71% of the time.

Figure 7: Mark Strength by Outcome



In an empirical study on mark strength, Lisa Ouellette observed that “courts often have difficulty applying these tests.”¹⁹⁴ According to her,

[t]he complex doctrine that has evolved around trademark strength and the likelihood of confusion appears to be a (largely unsuccessful) attempt to provide some analytical rigor to the essential questions of how strongly a mark identifies goods or services and how well it distinguishes those products from others in the marketplace.¹⁹⁵

Determining the bounds of an owner’s trademark requires more than just looking at the mark; it requires assessing what protection the trademark owner should be entitled to for that mark.¹⁹⁶ Distinctive marks are memorable as source indicators and possess greater conceptual strength to consumers.¹⁹⁷ Courts equate distinctiveness with a greater breadth of protection, are more willing to find confusing

194. Ouellette, *supra* note 40, at 353.

195. *Id.* at 360.

196. See *Virgin Enters. Ltd. v. Nawab*, 335 F.3d 141, 148 (2d Cir. 2003).

197. See *id.* (noting that consumers are more likely to attribute two products with more unique names to the same source versus two products with more generic names).

similarities,¹⁹⁸ and usually find that the strongest marks merit the widest range of protection.¹⁹⁹

Like the “black arts” of surveys, empirical studies confirm that courts judge mark strength intuitively.²⁰⁰ For instance, Beebe reported how courts failed to categorize the plaintiff’s mark in a specific category of distinctiveness in half of the cases he studied.²⁰¹ He observed that “considerations such as the comparative quality of the parties’ goods or the inherent distinctiveness of the plaintiff’s mark rarely aid in this inquiry.”²⁰² Others have variously criticized trademark strength as “needlessly open-ended”²⁰³ and “inconsistent.”²⁰⁴ One court acknowledged distinctiveness “is far from an exact science and that the differences between the classes, which is not always readily apparent . . . makes placing a mark in its proper context . . . tricky business at best.”²⁰⁵

As with survey evidence, Thomas McCarthy notes, that

a cynic would say that . . . when the court wants to find no infringement, it says that the average buyer is cautious and careful But if the judge thinks there is infringement, the judge sets the standard lower and says the average buyer is gullible and not so discerning.²⁰⁶

198. See, e.g., *First Sav. Bank, F.S.B. v. First Bank Sys., Inc.*, 101 F.3d 645, 655 (10th Cir. 1996) (“When the primary term is weakly protected to begin with, minor alterations may effectively negate any confusing similarity between the two marks.”).

199. See, e.g., *Ford Motor Co. v. Money Makers Auto. Surplus, Inc.*, No. 03CV493, 2005 WL 2464715, at *1, *3 (D. Neb. Sept. 14, 2005) (finding that the various Ford Motor Company marks at issue “are among the most famous marks in the world” and are “therefore entitled to the widest scope of protection”).

200. See, e.g., Thomas R. Lee, Eric D. DeRosia & Glenn L. Christensen, *Sophistication, Bridging the Gap, and the Likelihood of Confusion: An Empirical and Theoretical Analysis*, 98 TRADEMARK REP. 913, 913 (2008) (analyzing how courts rely on “precedent built on ‘personal intuition and subjective, internalized, stereotypes.’”); see also Beebe, *supra* note 6, at 1581 (describing the variation among circuits in their application of multifactor tests for likelihood of confusion).

201. Beebe, *supra* note 6, at 1633–35 (stating that some use of the spectrum was made in only 193 out of 331 cases and that the mark was placed in a specific category in only 164 cases).

202. *Id.* at 1645.

203. Timothy Denny Greene & Jeff Wilkerson, *Understanding Trademark Strength*, 16 STAN. TECH. L. REV. 535, 582 (2013).

204. Beebe, *supra* note 6, at 1633.

205. *Banff, Ltd. v. Federated Dep’t Stores, Inc.*, 841 F.2d 486, 489 (2d Cir. 1988).

206. 4 MCCARTHY, *supra* note 2, § 23:92; see also Ann Bartow, *Likelihood of Confusion*, 41 SAN DIEGO L. REV. 721, 747 (2004) (noting that judges give meaning to terms on a “case-by-case” basis).

The courts themselves regard likelihood of confusion merely “as a heuristic device to assist in determining whether confusion exists.”²⁰⁷ There is no requirement for the likelihood of confusion to consider survey evidence or mark strength. Eliminating both would simplify the likelihood of confusion and make it less prone to error.

4. *Consumer sophistication*

Consumer sophistication provides context to the consumer information available and the ability of consumers to discern between the marks.²⁰⁸ Courts analyze the degree of care reasonably expected of potential customers from the perspective of “the ordinary purchaser, buying under the normally prevalent conditions of the market and giving the attention such purchasers usually give in buying that class of goods.”²⁰⁹ More expensive products or services mean consumers take more time and effort when making decisions, and therefore, the likelihood of confusion decreases.²¹⁰ However, the defendant’s distribution methods may affect consumers’ degree of care, even when an individual product is not expensive.²¹¹

Scholars criticized the artificiality of consumer sophistication, likening it to expecting judges to perform a “Vulcan mind-meld” with consumers in the marketplace.²¹² Courts may easily project their normative view of how careful a consumer should be or their view of a

207. *Sullivan v. CBS Corp.*, 385 F.3d 772, 778 (7th Cir. 2004).

208. Andrew Martineau, *Imagined Consumers: How Judicial Assumptions About the American Consumer Impact Trademark Rights, for Better and for Worse*, 22 DEPAUL J. ART, TECH. & INTELL. PROP. L. 337, 352 (2012) (“This would seem to be a crucial part of the test, given that the standard for infringement is whether consumers are likely to be confused.”).

209. *Gen. Mills, Inc. v. Kellogg Co.*, 824 F.2d 622, 627 (8th Cir. 1987).

210. *Kibler v. Hall*, 843 F.3d 1068, 1080 (6th Cir. 2016) (observing that when consumers exercise caution in purchasing items, they are less likely to confuse their origins, such as “when consumers have expertise in the items and when the items are particularly expensive”).

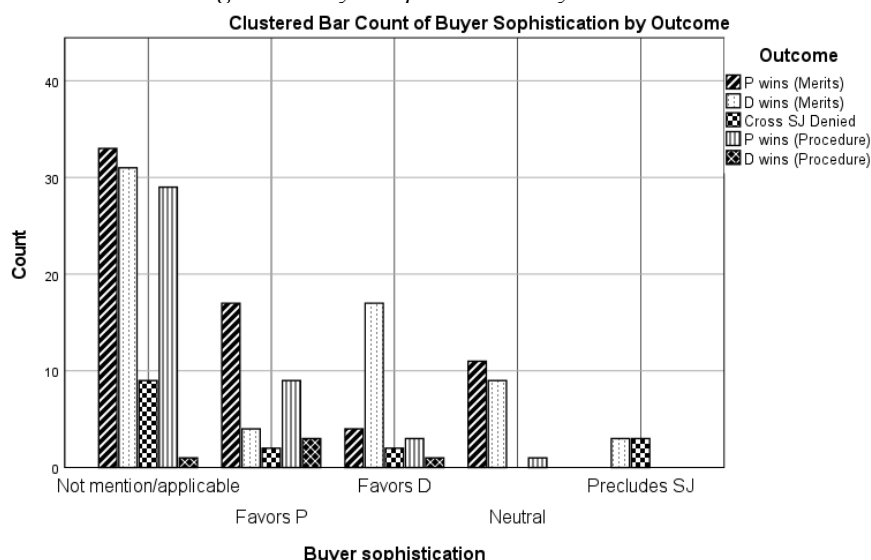
211. *See, e.g., ZW USA, Inc. v. PWD Sys., LLC*, 889 F.3d 441, 447–48 (8th Cir. 2018) (finding fact that parties sold their respective low-cost products on different websites under different trade names strongly cut against a likelihood of confusion).

212. *See* William E. Gallagher & Ronald C. Goodstein, *Inference Versus Speculation in Trademark Infringement Litigation: Abandoning the Fiction of the Vulcan Mind Meld*, 94 TRADEMARK REP. 1229, 1230 (2004) (criticizing the lack of empirical evidence required to validate an inference of likelihood of confusion, such as no requirement for consumer surveys or evidence showing actual confusion).

defendant's conduct.²¹³ But, like intent, surveys, and mark strength, consumer sophistication suffers from inherent capriciousness.

The dataset reveals that consumer sophistication comes up in 46% of the cases, among the lowest of all the *Polaroid* factors. In 18% of all cases, the courts favored plaintiffs on the consumer sophistication factor. When they did, plaintiffs won 49% of the time. In 14% of all cases, the courts favored defendants on the consumer sophistication factor. When this factor favored defendants, they won 63% of the time.

Figure 8: Buyer Sophistication by Outcome



Three irrelevant factors are plenty, but there is one final culprit. That is, the sheer multitude of factors courts must consider. The total number of factors makes the likelihood of confusion analysis difficult to deploy, bogging down courts and encouraging selective application. Instead, judges and juries rely on coherence-based reasoning to make sense of their findings to cope with the sheer number of factors.

D. Coherence-Based Reasoning

Over the past century, trademark law ossified the likelihood of confusion standard from pragmatic judge-made rules of thumb into a rigid and formalistic standard. The *Restatement (First) of Torts* merely

213. August Storck K.G. v. Nabisco, Inc., 59 F.3d 616, 618 (7th Cir. 1995) ("Many consumers are ignorant or inattentive, so some are bound to misunderstand no matter how careful a producer is.").

mentioned “the following factors are important,” and the early cases applied the factors loosely.²¹⁴ However, appeals courts chastised lower courts for failing to address each factor, with orders to reverse and remand.²¹⁵ We can deduce this formalism ended up burdening courts with an unwieldy craft, forcing judges to pay lip service to all the factors while systemically relying on only a few. At the same time, their opinions recite disclaimers that the likelihood of confusion factors act only as a guide and that no single factor is dispositive.

Studies show that experts do not integrate multifactor test (“MFT”) factors well.²¹⁶ Even using stringent tests to aid in decision-making can lead to consistent and predictable mistakes.²¹⁷ It may occur early in the decision-making process, and a single attribute can trigger coherence-based reasoning.²¹⁸

Trademark law expects courts to decipher between six and thirteen likelihood of confusion factors, which often point in opposite directions, yet still reach a coherent conclusion in every case.²¹⁹ Worse, the likelihood of confusion factors in each circuit are not exhaustive, with courts occasionally considering other factors such as geographical proximity.²²⁰

Courts are divided on whether “it is incumbent upon the district judge to engage in a deliberate review of each factor.”²²¹ Some emphatically state

214. RESTATEMENT (FIRST) OF TORTS § 729 (AM. L. INST. 1939).

215. Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581, 1593 (2006).

216. See, e.g., Robyn M. Dawes, *The Robust Beauty of Improper Linear Models in Decision Making*, 34 AM. PSYCH. 571, 573 (1979) (positing that experts in a field are better at selecting and coding information than integrating it).

217. See Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 779–80 (2001).

218. See Dan Simon, Daniel C. Krawczyk, & Keith J. Holyoak, *Construction of Preferences by Constraint Satisfaction*, 15 PSYCH. SCI. 331, 331 (2004) (suggesting that a single variable can initiate spreading coherence).

219. See *Tana v. Dantanna*, 611 F.3d 767, 775 n.7 (11th Cir. 2010) (explaining that this test “presupposes that various factors will point in opposing directions”; it is the job of the Court to determine the relative importance of the evidence probative of each factor in an effort to decide whether, “in light of the evidence as a whole, there is sufficient proof of a likelihood of confusion to warrant a trial of the issue”).

220. See *id.* at 781 (holding that “[t]he district court did not err in considering the geographic proximity of use as an eighth factor demonstrating the unlikelihood of confusion”); see also *J-B Weld Co. v. Gorilla Glue Co.*, 978 F.3d 778, 789 (11th Cir. 2020) (“While all seven factors must be considered, they are not necessarily exhaustive if other evidence is probative of a likelihood of confusion.”).

221. Compare *Arrow Fastener Co. v. Stanley Works*, 59 F.3d 384, 400 (2d Cir. 1995) (discussing each factor), with *Bumble Bee Seafoods LLC v. UFS Indus., Inc.*, No. 04

that “the factors are not truly independent—depending on the context, a strong showing as to one factor may serve to make a different factor more or less important.”²²² Yet others rule only on a few key factors, allowing them to resolve the dispute without needing a trial.²²³

Without meaningful guidance, courts weigh those factors impressionistically. Beebe’s study confirms that judges in the likelihood of confusion cases employ “‘fast and frugal’ heuristics to short-circuit the multifactor analysis.”²²⁴ Coherence-based reasoning operates bidirectionally to fit together how a judge decides the factors,²²⁵ both preceding the decision and in forming its basis.²²⁶ In other words, fact-finders assessing a likelihood of confusion test will look at the evidence as non-independently relative to the final decision.²²⁷ Consequently, the resulting decision is biased because, as Dan Simon explains, “the hard case morphs into an easy one” in the mind of the fact-finder.²²⁸

Formulating optimal legal rules requires judges to balance factors while taking account of “false positive” errors (i.e., prohibiting beneficial conduct) versus “false negative” errors (i.e., permitting harmful conduct).²²⁹ This task requires judges to access information on the frequency and impact of the error, the likelihood of deterrence,

Civ. 2105, 2004 WL 1637017, at *7 (S.D.N.Y. July 20, 2004) (discussing only the relevant factors).

222. *Select Comfort Corp. v. Baxter*, 996 F.3d 925, 934 (8th Cir. 2021).

223. *See, e.g., Hamilton Int’l Ltd. v. Vortic LLC*, 486 F. Supp. 3d 657, 666 (S.D.N.Y. 2020) (“As the Court noted in its summary judgment opinion, a number of the *Polaroid* factors are not helpful to this case.”), *aff’d*, 13 F.4th 264 (2d Cir. 2021).

224. Beebe, *supra* note 6, at 1581; Tierney, *supra* note 150, at 235–36 (“[M]uch of the time spent going through the list of factors in any given case is in reality just an attempt to justify a predetermined conclusion about the likelihood of confusion . . .”).

225. Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 515–16 (2004) [hereinafter Simon, *Third View of the Black Box*].

226. *See, e.g., Dan Simon et al., The Redux of Cognitive Consistency Theories: Evidence Judgments by Constraint Satisfaction*, 86 J. PERSONALITY & SOC. PSYCH. 814, 816 (2004).

227. Dan Simon, *The Limited Diagnosticity of Criminal Trials*, 64 VAND. L. REV. 143, 195 (2011).

228. Simon, *Third View of the Black Box*, *supra* note 225, at 517 (describing studies where coherence-based reasoning caused subjects who found for the defendant and those who found for the plaintiff to be more confident the evidence supported their view after they had issued their verdict).

229. Andrew I. Gavil & Steven C. Salop, *Probability, Presumptions and Evidentiary Burdens in Antitrust Analysis: Revitalizing the Rule of Reason for Exclusionary Conduct*, 168 U. PA. L. REV. 2107, 2119 (2020).

and the cost to the administrative process.²³⁰ On occasion, courts themselves express frustration with the likelihood of confusion tests, acknowledging that “[a]lthough our test for a likelihood of confusion is well-developed, some uncertainty remains as to *when* confusion must exist in order to support a trademark infringement claim.”²³¹

The takeaway is that an overload of factors demands too much from judges and forces them to stampede over those they deem less significant. In the absence of direct evidence of confusion, courts must ascertain it through a host of proxy factors.²³² Under these trying circumstances, Beebe empirically observed intent and actual confusion playing an outsized role in coloring how courts treated the other likelihood of confusion factors, confirming their perniciousness.²³³

As Michael Grynberg noted, “[e]ven if judges do no more than applying heuristics of questionable quality to the disposition of trademark claims, channeling the process through a consistent framework aids litigants in identifying and accommodating the factors that guide fact finding.”²³⁴ The question then is, how many factors should we retain?

Only a few, argued Beebe, pointing out that judges in the likelihood of confusion cases find only a few factors probative anyway.²³⁵ Indeed, cases in the dataset recognize that courts can short circuit the process and focus on just a few factors.²³⁶ Beebe recommended three or four

230. See *id.* at 2119–20 (asserting that formulating an optimal legal standard involves considering error costs, deterrence, and administrative costs).

231. See, e.g., *Select Comfort Corp. v. Baxter*, 996 F.3d 925, 934 (8th Cir. 2021).

232. Laura A. Heymann, *The Reasonable Person in Trademark Law*, 52 ST. LOUIS U. L.J. 781, 783 (2008).

233. Beebe, *supra* note 6, at 1620–21 (“Intent and, to a lesser degree, actual confusion appear to exert such a coherence-shifting influence when they favor a likelihood of confusion. Indeed, in the forty-nine opinions in which both findings were made, thirty-four (69%) of them found that all the factors favored a likelihood of confusion.”).

234. Michael Grynberg, *The Judicial Role in Trademark Law*, 52 B.C. L. REV. 1283, 1305 (2011) [hereinafter Grynberg, *Judicial Role in Trademark Law*].

235. Beebe, *supra* note 6, at 1614 (“Like any human decision makers, district judges attempt to decide both efficiently and accurately. In pursuit of efficiency, they consider only a few factors. In pursuit of accuracy, they consider the most decisive factors.”).

236. *Eclipse Assocs. Ltd. v. Data Gen. Corp.*, 894 F.2d 1114, 1118 (9th Cir. 1990) (“These tests were not meant to be requirements or hoops that a district court need jump through to make the determination.”); see also *R.H. Donnelley Inc. v. USA Northland Directories, Inc.*, No. Civ.04-4144, 2004 WL 2713248, at *6 (D. Minn. Nov. 19, 2004) (folding similarity and intent); *Ironhawk Techs. v. Dropbox, Inc.*, 994 F.3d 1107, 1123–24 (9th Cir. 2021) (folding actual confusion and sophistication); *CDOC*,

“core factors” informing “consumer perception in the marketplace rather than judicial perception in the courtroom.”²³⁷ Alejandro Mejías went further, recommending just two—similarity of marks and proximity of goods, as “adding any other relevant factors, instead of using unmanageable and misleading large lists of factors that are extremely difficult to balance, seems to be more in line with the thesis of scientific research on decision-making.”²³⁸ The next Part explains why actual confusion, mark similarity, and competitive proximity should form the core factors and why these factors, together with fair use safe harbors for expressive and descriptive uses, should form the rules of thumb in trademark law.

II. RULES OF THUMB

Mark similarity, goods and services, and evidence of actual confusion anchor the likelihood of confusion test as the most relevant factors.²³⁹ Jurisprudence supports that view. In one case from the dataset, the Ninth Circuit has described a “trinity [that] constitutes the most crucial body of the *Sleekcraft* analysis”—mark similarity, goods/services similarity, and marketing and advertising channels.²⁴⁰ Safe harbors protect core policies most in danger of being invaded by trademark expansionism while making it simpler and cheaper for businesses to do their due diligence and comply with the law.²⁴¹ This Part explains why.

A. Actual Confusion

Actual confusion is the most direct and decisive evidence of confusion.²⁴² Courts explain that where confusion occurred, it “is of

Inc. v. Liberty Bankers Life Ins. Co., 844 F. App’x 357, 362 (Fed. Cir. 2021) (folding competitive proximity and sophistication).

237. Beebe, *supra* note 6, at 1646.

238. Mejías, *supra* note 106, at 348 (concentrating the analysis on the main two factors).

239. See *infra* Part III.

240. GoTo.com, Inc. v. Walt Disney Co., 202 F.3d 1199, 1207 (9th Cir. 2000); see also Stone Creek, Inc. v. Omnia Italian Design, Inc., 862 F.3d 1131, 1136 (9th Cir. 2017) (“Two particularly probative factors are the similarity of the marks and the proximity of the goods.” (citing *Lindy Pen Co. v. Bic Pen Corp.*, 796 F.2d 254, 256–57 (9th Cir. 1986))).

241. See *infra* Part II.

242. Beebe, *supra* note 6, at 1608 (finding a 92% plaintiff success rate where the court found actual confusion); *Tana v. Dantanna’s*, 611 F.3d 767, 779 (11th Cir. 2010) (“The last factor, actual confusion in the consuming public, is the most persuasive evidence in assessing likelihood of confusion.”); see also *Groeneveld Transp. Efficiency*,

course convincing evidence that confusion is likely to occur.”²⁴³ As a policy lever, actual confusion gives courts the ability to anchor their analysis in real-world characteristics. In addition, the evidence is pre-existing, does not depend on the vagaries of survey design, and should make it easier for courts to dispose of cases pretrial.²⁴⁴

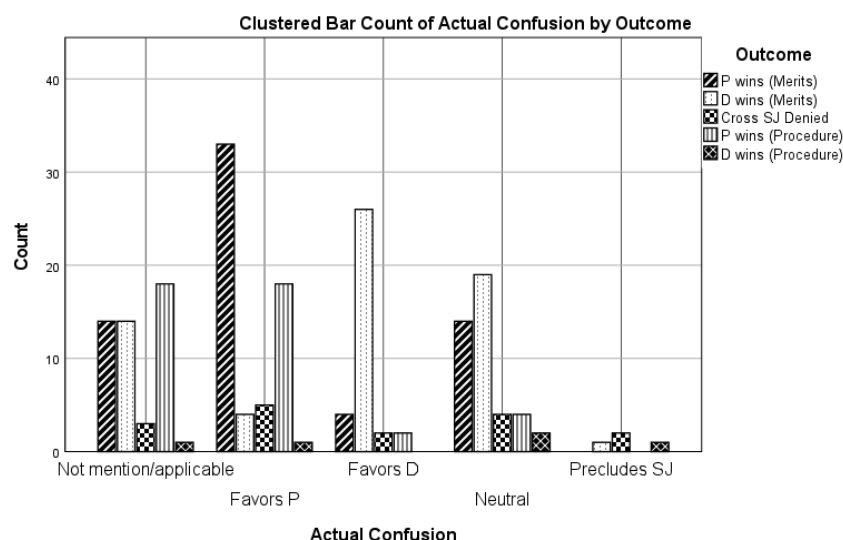
The dataset reveals that actual confusion comes up in 74% of the cases. In 32% of all cases, the courts favored plaintiffs on the actual confusion factor. When they did, plaintiffs won 54% of the time. In 18% of all cases, the courts favored defendants on the consumer sophistication factor. When the court favored defendants, they won 77% of the time.

Inc. v. Lubecore Int’l, Inc., 730 F.3d 494, 517 (6th Cir. 2013) (“Nothing shows the likelihood of confusion more than the fact of actual confusion.”); Variety Stores, Inc. v. Walmart Inc., 852 F. App’x 711, 720 (4th Cir. 2021) (“[A]ctual confusion, is the ‘most important factor’”); John Benton Russell, *New Tenth Circuit’s Standards: Competitive Keyword Advertising and Initial Interest Confusion in 1-800 Contacts v. Lens.com*, 30 BERKELEY TECH. L.J. 993, 1000 (2015) (“[C]ourts across several circuits view this as the strongest evidence a plaintiff can present in a trademark infringement case.”); Mark D. Robins, *Actual Confusion in Trademark Infringement Litigation: Restraining Subjectivity Through a Factor-Based Approach to Valuing Evidence*, 2 NW. J. TECH. & INTELL. PROP. 117, 117 (2004) (“In a case where all other circumstances point to a finding of non-infringement, significant evidence of actual confusion dramatically alters the equation.”).

243. *Morningside Grp. Ltd. v. Morningside Cap. Grp., L.L.C.*, 182 F.3d 133, 141 (2d Cir. 1999).

244. I am grateful to Jon Lee for this insight.

Figure 9: Actual Confusion by Outcome



At the bottom, the inquiry concerns whether there was confusion that could lead to “a diversion of sales, damage to goodwill, or loss of control over reputation.”²⁴⁵ For that reason, courts look for actual confusion among “prospective purchasers of [plaintiff’s] products.”²⁴⁶ Relevant circumstances include the extent of the parties’ advertising, the length of time the allegedly infringing product has been advertised, or any other factor that might influence the likelihood that actual confusion would be reported.²⁴⁷

Courts accept both anecdotal and survey evidence indicating actual confusion.²⁴⁸ This Article explained in Section I.C.2 that courts should avoid survey evidence in its current manifestation. As to anecdotal evidence, there is no absolute number of instances of actual confusion that must be met to win in a likelihood of confusion analysis. Rather, courts look to the totality of the circumstances in evaluating the evidence of actual confusion.²⁴⁹ For example, “[i]nquiries about the relationship between an owner of a mark and an alleged infringer do

245. *Reply All Corp. v. Gimlet Media, LLC*, 843 F. App’x 392, 397 (2d Cir. 2021) (quoting *Lang v. Ret. Living Publ’g Co.*, 949 F.2d 576, 583 (2d Cir. 1991)).

246. *Lang*, 949 F.2d at 583; *SLY Mag., LLC v. Weider Publ’ns L.L.C.*, 529 F. Supp. 2d 425, 441 (S.D.N.Y. 2007), *aff’d*, 346 F. App’x 721 (2d Cir. 2009).

247. *Jellibeans, Inc. v. Skating Clubs Ga., Inc.*, 716 F.2d 833, 844 (11th Cir. 1983).

248. *George & Co. v. Imagination Ent. Ltd.*, 575 F.3d 383, 398 (4th Cir. 2009).

249. *AmBrit, Inc. v. Kraft, Inc.*, 812 F.2d 1531, 1543 (11th Cir. 1986).

not amount to actual confusion.”²⁵⁰ Testimony from one customer—the mark’s owner, and its employee—arguing customers mistakenly visited the defendant’s store when intending to visit the owner’s store constitutes *de minimis* evidence of actual confusion.²⁵¹ While “[i]solated instances of [actual] confusion are insufficient to support a finding of likely confusion,”²⁵² courts have found confusion by five people,²⁵³ or even one person increases the likelihood of confusion.²⁵⁴ At the same time, “it is well established that no actual confusion is required to prove a case of trademark infringement.”²⁵⁵ Courts have justified this conclusion “[b]ecause of the difficulty in garnering such evidence.”²⁵⁶

Confusion must be by the “actual consuming public” and therefore anchored in a real-world context.²⁵⁷ The absence of actual confusion “over a substantial period . . . creates a strong inference that there is no likelihood of confusion.”²⁵⁸ In combining the two ideas, “[s]hort-lived confusion or confusion of individuals casually acquainted with a business is worthy of little weight, while confusion of actual customers of a business is worthy of substantial weight.”²⁵⁹ Where a large volume of contacts or transactions could give rise to confusion, and only limited instances of confusion present themselves, courts give evidence of actual confusion little weight.²⁶⁰

250. *Reply All Corp.*, 843 F. App’x at 398 (alteration in original).

251. *RVC Floor Decor, Ltd. v. Floor & Decor Outlets of Am., Inc.*, 527 F. Supp. 3d 305, 326 (E.D.N.Y. 2021).

252. *Progressive Distrib. Servs., Inc. v. United Parcel Serv., Inc.*, 856 F.3d 416, 433 (6th Cir. 2017).

253. *AWGI, LLC v. Atlas Trucking Co.*, 998 F.3d 258, 267 (6th Cir. 2021).

254. *Innovation Ventures, LLC v. N2G Distrib., Inc.*, 763 F.3d 524, 536 (6th Cir. 2014) (“[A] single instance of actual confusion can, in some cases, ‘increase the likelihood of confusion.’”).

255. *Louis Vuitton Malletier S.A. v. Haute Diggity Dog, LLC*, 507 F.3d 252, 263 (4th Cir. 2007).

256. *Monster Energy Co. v. BeastUp LLC*, 395 F. Supp. 3d 1334, 1358 (E.D. Cal. 2019) (alteration in original).

257. *Rearden LLC v. Rearden Com., Inc.*, 683 F.3d 1190, 1210 (9th Cir. 2012).

258. *CareFirst of Md., Inc. v. First Care, P.C.*, 434 F.3d 263, 269 (4th Cir. 2006) (finding an inference of no likelihood of confusion where there was no evidence of confusion for nine years).

259. *Safeway Stores, Inc. v. Safeway Disc. Drugs, Inc.*, 675 F.2d 1160, 1167 (11th Cir. 1982).

260. *George & Co. v. Imagination Ent. Ltd.*, 575 F.3d 383, 399 (4th Cir. 2009) (explaining that “the company’s failure to uncover more than a few instances of actual confusion creates a presumption against likelihood of confusion in the future” when there are so many opportunities for confusion to occur).

Some courts hold that a lack of evidence of actual confusion does not create a presumption of no confusion but is “simply a factor in the court’s analysis.”²⁶¹ However, as a policy lever, it gives courts the ability to anchor their analysis in real-world characteristics. In addition, the evidence is pre-existing, does not depend on the vagaries of survey design, and should make it easier for courts to dispose of cases pretrial.²⁶² If found, it is worth its weight in gold, tipping the balance in the plaintiff’s favor more than any other factor.²⁶³

B. Mark Similarity

Three axioms apply to the “similarity” analysis: (1) marks should be considered in their entirety and as they appear in the marketplace; (2) similarity is judged by appearance, sound, and meaning; (3) and, similarities weigh more heavily than differences.²⁶⁴ Courts determine whether a mark confuses the public when viewed alone to account for the possibility that similar marks “may confuse consumers who do not have both marks before them but who may have a general, vague, or even hazy, impression or recollection of the other party’s mark.”²⁶⁵

At the most basic level, marks are confusingly similar if “ordinary consumers would likely conclude that [the two products] share a common source, affiliation, connection or sponsorship.”²⁶⁶ Identical, even dominant, features do not “automatically mean that two marks are similar.”²⁶⁷ Courts look to “the overall impression created by the marks, not merely compare individual features,” and “may consider the marks’ visual, aural, and definitional attributes and compare the trade dress of the products in determining whether the total effect conveyed by the two marks is confusingly similar.”²⁶⁸

261. See, e.g., *Am. Dairy Queen Corp. v. W.B. Mason Co.*, 543 F. Supp. 3d 695, 718–19 (D. Minn. 2021).

262. I am grateful to Jon Lee for this insight.

263. “Bridging the gap” reported a 71% win rate but given its relative infrequency (25% versus 74% for actual) and large overlap with competitive proximity (which could explain why it is even at 25%, for that matter 71%), the better view is to discount it. *Lim*, *supra* note 41.

264. *Entrepreneur Media, Inc. v. Smith*, 279 F.3d 1135, 1144 (9th Cir. 2002).

265. *Maker’s Mark Distillery, Inc. v. Diageo N. Am., Inc.*, 679 F.3d 410, 421 (6th Cir. 2012).

266. *Fisons Horticulture, Inc. v. Vigoro Indus., Inc.*, 30 F.3d 466, 477 (3d Cir. 1994).

267. *Sensient Techs. Corp. v. SensoryEffects Flavor Co.*, 613 F.3d 754, 764 (8th Cir. 2010) (quoting *Gen. Mills, Inc. v. Kellogg Co.*, 824 F.2d 622, 627 (8th Cir. 1987)).

268. *Luigino’s, Inc. v. Stouffer Corp.*, 170 F.3d 827, 830 (8th Cir. 1999); see also *Leelanau Wine Cellars, Ltd. v. Black & Red, Inc.*, 502 F.3d 504, 516–17 (6th Cir. 2007) (explaining that courts give similarity considerable weight and consider the

As the Eleventh Circuit noted, “[i]f a trademark operates in a crowded field of similar marks on similar goods or services, slight differences in names may be meaningful because consumers will not likely be confused between any two of the crowd and may have learned to carefully pick out one from the other.”²⁶⁹ Similarly, mark similarity takes prominence when the goods are direct competitors in the marketplace.²⁷⁰ The goods or services will likely be virtually identical. However, the marks need not be as similar for there to be a likelihood of confusion.²⁷¹

Courts extol the importance of mark similarity.²⁷² Beebe found it was “by far the most important factor.”²⁷³ In injunction cases, 83% of plaintiffs who won the similarity factor prevailed in the likelihood of confusion analysis, as did 90% in plaintiff summary judgment motions.²⁷⁴ The dataset reveals that only 23% of plaintiffs in injunction cases who won the similarity factor prevailed in likelihood of confusion analysis, a far lower number than before. However, 90% in plaintiff summary judgment motions succeeded just as before.

Fifteen years later, the dataset also reveals that mark similarity still comes up in 85% of the cases, the most frequently invoked factor of them all. In 61% of all cases, the courts favored plaintiffs on the mark similarity factor. When they did, plaintiffs won 47% of the time on the merits. In 18% of all cases, the courts favored defendants on the mark similarity factor. When the court favored defendants, they won 88% of the time on the merits.

“pronunciation, appearance, and verbal translation of conflicting marks,” and courts must “view marks in their entirety and focus on their overall impressions, not individual features” (quoting *AutoZone, Inc. v. Tandy Corp.*, 373 F.3d 786, 795–96 (6th Cir. 2004))).

269. *Fla. Int’l Univ. Bd. Trs. v. Fla. Nat’l Univ., Inc.*, 830 F.3d 1242, 1260 (11th Cir. 2016) (internal quotation marks omitted).

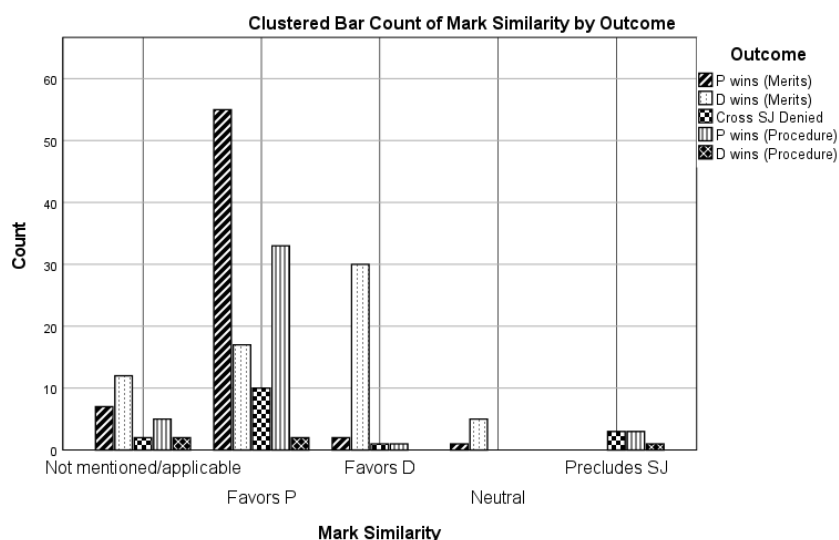
270. *Bos. Duck Tours, LP v. Super Duck Tours, LLC*, 531 F.3d 1, 30 (1st Cir. 2008).

271. *Bridgestone Ams. Tire Operations, LLC v. Fed. Corp.*, 673 F.3d 1330, 1337 (Fed. Cir. 2012) (likelihood of confusion between “Potenza” and “Turanza” marks was greater because both referred to tires).

272. *Equitable Nat’l Life Ins. Co. v. AXA Equitable Life Ins. Co.*, 434 F. Supp. 3d 1227, 1246 (D. Utah 2020) (“The similarity of the marks is the ‘first and most important factor.’”); *Juul Labs, Inc. v. 4X PODS.*, 439 F. Supp. 3d 341, 353 (D.N.J. 2020) (calling it “[t]he single most important factor in determining likelihood of confusion”); *GoTo.com, Inc. v. Walt Disney Co.*, 202 F.3d 1199, 1205 (9th Cir. 2000) (noting that mark similarity “has always been considered a critical question in the likelihood-of-confusion analysis”).

273. Beebe, *supra* note 6, at 1623.

274. *Id.* at 1625.

Figure 10: Mark Similarity by Outcome

The dataset reveals interesting dynamics between rivalry and mark similarity. When the parties' marks were similar and the parties were rivals, plaintiffs won 47% of the time. Similarly, when the parties were non-rivals, plaintiffs also won 47% of the time. However, there is a difference in how often the defendant wins. Defendants only win 10% of the time when they are rivals and 24% when they are not rivals, underscoring the expectedly powerful role rivalry plays, but in an asymmetrical way.

One possible explanation is that similarity between the marks makes consumers more likely to become confused about the source. Extremely similar marks or goods may suggest counterfeiting and free riding. Parodies, comparative advertising, and nominative use make consumers less likely to be confused, even if the third party uses the identical term.

Courts even dispensed entirely with the likelihood of confusion test when parties' marks were identical, a conclusion with implications for early off-ramping parties, as discussed in Part III.²⁷⁵ Where a defendant uses a counterfeit mark, such use is deemed inherently confusing to a

275. See *Phillip Morris USA Inc. v. Shalabi*, 352 F. Supp. 2d 1067, 1073 (C.D. Cal. 2004) (“[I]n cases involving counterfeit marks, it is unnecessary to perform the step-by-step examination . . . because counterfeit marks are inherently confusing.”); *Daimler AG v. A-Z Wheels LLC*, 334 F. Supp. 3d 1087, 1096 (S.D. Cal. 2018) (“It is not necessary for the Court to analyze the likelihood of confusion test here considering Defendants’ use the identical MERCEDES-BENZ mark.”).

customer.²⁷⁶ As McCarthy explained “[c]ases where a defendant uses an identical mark on competitive goods hardly ever find their way into the appellate reports. Such cases are ‘open and shut’ and do not involve protracted litigation to determine liability for trademark infringement.”²⁷⁷ This is because “confusing the customer is the whole purpose of creating counterfeit goods.”²⁷⁸ Such cases create a presumption of harm such that the factor may stampede the likelihood of confusion analysis entirely.²⁷⁹

Aside from the simplest forms of counterfeiting, the threshold triggering confusion, and more so *likely* confusion, exists only as a relative measure where reasonable minds may differ. Unlike real property, there are no metes and bounds. This lack of boundaries presents interpretive challenges that Michael Grynberg and Graeme Austin independently attributed to the likelihood of confusion’s current uncertainty.²⁸⁰ The problem is common to other areas of the law as well. For instance, copyright law’s substantial similarity standard suffers many of the same ills as the likelihood of confusion and demands reconsideration.²⁸¹

276. *EAT BBQ LLC v. Walters*, 47 F. Supp. 3d 521, 530 (E.D. Ky. 2014) (“[T]here is almost never a dispute regarding confusion.”).

277. *Wynn Oil Co. v. Thomas*, 839 F.2d 1183, 1191 (6th Cir. 1988) (quoting 4 MCCARTHY, *supra* note 2, § 23:20).

278. *Gucci Am., Inc. v. Duty Free Apparel, Ltd.*, 286 F. Supp. 2d 284, 287 (S.D.N.Y. 2003).

279. *See* *Threeline Imps., Inc. v. Vernikov*, 239 F. Supp. 3d 542, 561 (E.D.N.Y. 2017); *see also* *Dish Network L.L.C. v. Siddiqi*, No. 18 CV 4397, 2019 WL 5781945, at *4 (S.D.N.Y. 2019) (“[I]t is not necessary to perform the step-by-step examination of each *Polaroid* factor’ when a counterfeit mark is at issue.”); *Halo Optical Prods., Inc. v. Liberty Sport, Inc.*, No. 14-cv-00282, 2017 WL 1082443, at *11 (N.D.N.Y. Mar. 22, 2017) (“[W]hen dealing with an identical mark . . . courts are not necessarily required to analyze the *Polaroid* factors.”); *Gucci Am.*, 286 F. Supp. 2d at 287 (“[T]he Court need not undertake a factor-by-factor analysis under *Polaroid* because counterfeits, by their very nature, cause confusion.”).

280. Grynberg, *Judicial Role in Trademark Law*, *supra* note 234, at 1303 (“Trademark’s fundamental inquiry, whether a likelihood of confusion exists, invites judicial lawmaking in no small part because the term ‘likelihood of confusion’ presents an interpretive problem.”); Graeme W. Austin, *Tolerating Confusion About Confusion: Trademark Policies and Fair Use*, 50 ARIZ. L. REV. 157, 160 (2008) (“There is considerable uncertainty about some of the key questions that are germane to the factual inquiry at the heart of the likelihood of confusion analysis.”).

281. *See, e.g.*, Daryl Lim, *Saving Substantial Similarity*, 73 FLA. L. REV. 591, 640–41 (2021) [hereinafter Lim, *Saving Substantial Similarity*] (explaining how the substantial similarity standard generates “capricious and wrong results”).

Courts use sights, sounds, and meaning to make snap judgments about mark similarity.²⁸² These heuristics allow judges to rely on “a small set of cheap and reliable factors that are close enough to the ideal.”²⁸³ Adam Samaha approves of it since “[p]rioritizing the judge’s impressions about the similarity of marks, therefore, tends toward the high values of trademark law at bargain basement prices.”²⁸⁴ Defendants can easily compare visual or aural elements in context, making this a useful factor to encourage due diligence.²⁸⁵

The key takeaway is that the commercial context matters.²⁸⁶ Marks should not be compared side-by-side as they might be shown in the courtroom.²⁸⁷ Instead, courts determine whether the public would confuse the marks when viewed alone because some highly similar marks can confuse consumers that view them without appropriate commercial context.²⁸⁸ For this reason, courts cannot dissect marks since consumers encounter them in their entirety in those settings. Instead, courts focus on their overall impressions rather than on their features.²⁸⁹ That which qualifies as mark similarity also disqualifies intent, survey evidence, mark strength, and consumer sophistication.

C. *Competitive Proximity*

Competitive proximity tells courts how likely consumers are to assume an association between the marks used on related products.²⁹⁰ For example, “[t]he similarities between the parties’ distribution channels and marketing strategies suggest an overlapping general class

282. Adam M. Samaha, *Looking Over a Crowd—Do More Interpretive Sources Mean More Discretion?*, 92 N.Y.U. L. REV. 554, 614 (2017) (“[A]ccurately estimating the probability of consumer confusion can require a snap judgment, which often is how consumers actually formulate impressions and make purchasing decisions.”).

283. *Id.*

284. *Id.*

285. See 4 MCCARTHY, *supra* note 2, § 23:21 (discussing the “sound, sight, and meaning” test for mark similarity).

286. Flower Mfg., LLC v. CareCo, LLC, 466 F. Supp. 3d 797, 814 (N.D. Ohio 2020) (“I am to judge the marks’ similarity as they appear in their commercial context.”); see also Homeowners Grp., Inc. v. Home Mktg. Specialists, Inc., 931 F.2d 1100, 1109 (6th Cir. 1991) (“[A] court must determine, in the light of what occurs in the marketplace, whether the mark will be confusing to the public when singly presented.” (internal quotation marks omitted) (quoting Wynn Oil Co. v. Thomas, 839 F.2d 1183, 1187 (6th Cir. 1988))).

287. Homeowners Grp., Inc., 931 F.2d at 1106.

288. Daddy’s Junky Music Stores, Inc. v. Big Daddy’s Fam. Music Ctr., 109 F.3d 275, 283 (6th Cir. 1997) (quoting Wynn Oil, 839 F.2d at 1188).

289. AutoZone, Inc. v. Tandy Corp., 373 F.3d 786, 795 (6th Cir. 2004).

290. Virgin Enters. Ltd. v. Nawab, 335 F.3d 141, 150 (2d Cir. 2003).

of consumers of the parties' products."²⁹¹ However, two products or services within the same general field do not automatically trigger a likelihood of confusion.²⁹² Similarly, a high percentage of overlap in "an extremely small subset of products does not demonstrate a high degree of relatedness."²⁹³

Services and goods within the same broad industry are not necessarily related. Rather, related services are marketed and consumed such that buyers are likely to believe that the services come from a common company.²⁹⁴ Courts examine "how and to whom the respective goods or services of the parties are sold."²⁹⁵ Less likelihood of confusion exists where the goods are sold through different avenues, "parties have different customers[,] and [they] market their goods or services in different ways."²⁹⁶ "[I]f the parties compete directly, confusion is likely" between sufficiently similar marks.²⁹⁷ "[I]f the goods and services are somewhat related, but not competitive, then the likelihood of confusion will turn on other factors[.]"²⁹⁸ [F]inally, if the products are unrelated, confusion is highly unlikely."²⁹⁹

The dataset reveals that competitive proximity comes up in 74% of the cases. In 51% of all cases, the courts favored plaintiffs on the competitive proximity factor. When they did, plaintiffs won 45% of the time. In 14% of all cases, the courts favored defendants on the competitive proximity factor. When the court favored defendants, they won 85% of the time.

291. *Monster Energy Co. v. BeastUp LLC*, 395 F. Supp. 3d 1334, 1359 (E.D. Cal. 2019). Other circuits use similar formulations. *See e.g.*, *Therma-Scan, Inc. v. Thermoscan, Inc.*, 295 F.3d 623, 632 (6th Cir. 2002) (noting that direct rivalry through similar goods or services is likely confusing).

292. *Matrix Motor Co. v. Toyota Jidosha Kabushiki Kaisha*, 290 F. Supp. 2d 1083, 1092 (C.D. Cal. 2003).

293. *AutoZone, Inc.*, 373 F.3d at 798 ("[I]f [the defendant] stocked only five types of batteries all of which were also sold by [the plaintiff], the overlap would be 100%, even though in reality [the defendant] and [the plaintiff] would share only five products of the approximately 55,000 offered by [the plaintiff].").

294. *Daddy's Junky Music Stores, Inc. v. Big Daddy's Fam. Music Ctr.*, 109 F.3d 275, 282–83 (6th Cir. 1997) (quoting *Homeowners Grp., Inc. v. Home Mktg. Specialists, Inc.*, 931 F.2d 1100, 1109 (6th Cir. 1991)).

295. *Leelanau Wine Cellars, Ltd. v. Black & Red, Inc.*, 502 F.3d 504, 519 (6th Cir. 2007).

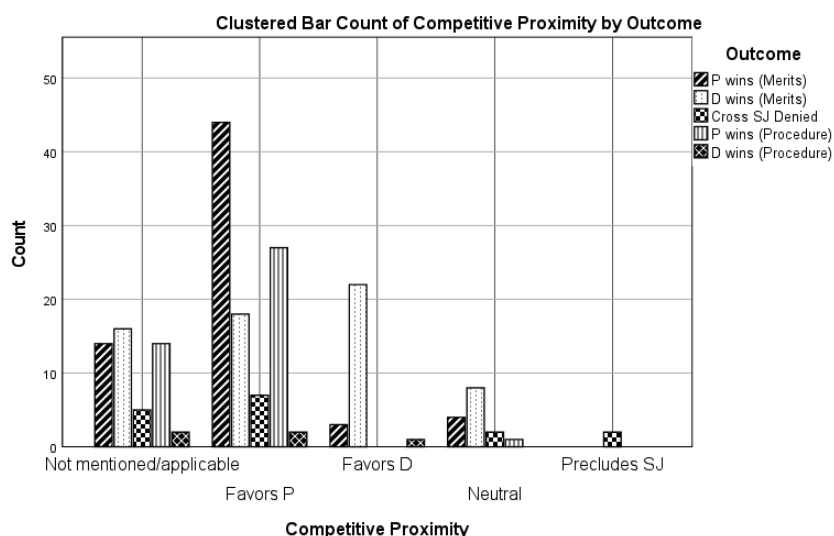
296. *Therma-Scan*, 295 F.3d at 636.

297. *Kellogg Co. v. Toucan Golf, Inc.*, 337 F.3d 616, 624 (6th Cir. 2003).

298. *Id.*

299. *Id.*

Figure 11: Competitive Proximity by Outcome



Competitive proximity encompasses adjacent *Polaroid* factors. One example is the likelihood plaintiffs or defendants will expand into each other's market or the "bridging the gap" factor.³⁰⁰ The likelihood that consumers will confuse the sources of parties' products increases when there is a "strong possibility that either party will expand its business to compete with the other's."³⁰¹ This confusion may happen when goods and services are complementary, sold to the same class of purchasers, or similar in use and function.³⁰²

Courts examine the two concepts in tandem with each other.³⁰³ Consider *Kohler Co. v. Bold International FZCO*,³⁰⁴ where the court noted that "[b]ridging the gap" refers to the likelihood that the senior

300. *Disney Enters., Inc. v. Sarelli*, 322 F. Supp. 3d 413, 434 (S.D.N.Y. 2018) ("The third and fourth *Polaroid* factors, respectively, address the proximity of the goods or services at issue and the possibility that the senior user will 'bridge the gap,' or expand the scope of its business and enter the market of the junior user. Thus, these two distinct but related factors 'focus on the degree to which the [parties'] products currently compete with each other or are likely to compete with each other in the future.'" (alteration in original) (quoting *Medici Classics Prods., LLC v. Medici Grp., LLC*, 683 F. Supp. 2d 304, 311–12 (S.D.N.Y. 2010))).

301. *Kibler v. Hall*, 843 F.3d 1068, 1082 (6th Cir. 2016).

302. *AMF Inc. v. Sleekcraft Boats*, 599 F.2d 341, 350 (9th Cir. 1979).

303. *RVC Floor Decor, Ltd. v. Floor & Decor Outlets Am., Inc.*, 527 F. Supp. 3d 305, 325 (E.D.N.Y. 2021) ("For the same reasons explained [in] the 'competitive proximity' analysis, the parties serve the same market and any gap has already been bridged.").

304. 422 F. Supp. 3d 681 (E.D.N.Y. 2018).

user . . . will enter into the same market as that of the junior user . . . where the goods are not yet in close competitive proximity.”³⁰⁵ When the parties’ goods are the same, courts simply fold this factor into competitive proximity as there is no gap to bridge.³⁰⁶ In this case, a consumer seeing the goods or services would likely be confused about their source.³⁰⁷

Another example is the degree of care the consumer might exercise in purchasing the parties’ goods, as mentioned in Section I.C.4. Courts look both to the “relative sophistication of the relevant consumer”³⁰⁸ and the cost of the item³⁰⁹ in determining the degree of care likely to be exercised by the purchaser. The “reasonably prudent consumer” is expected “to be more discerning—and less easily confused—when [they are] purchasing expensive items.”³¹⁰ Conversely, customers may be less careful when purchasing inexpensive products, thus making confusion more likely.³¹¹

“Bridging the gap” rarely arose, only in 25% of the cases. In 7% of all cases, the courts favored plaintiffs on the “bridging the gap” factor. When they did, plaintiffs won 71% of the time. In 6% of all cases, the courts favored defendants on the “bridging the gap” factor. When the court favored defendants, they won 83% of the time.

305. *Id.* at 725.

306. *Star Indus., Inc. v. Bacardi & Co.*, 412 F.3d 373, 387 (2d Cir. 2005) (“Because . . . [the parties’] products are already in competitive proximity, there is really no gap to bridge, and this factor is irrelevant to the *Polaroid* analysis in this case.”).

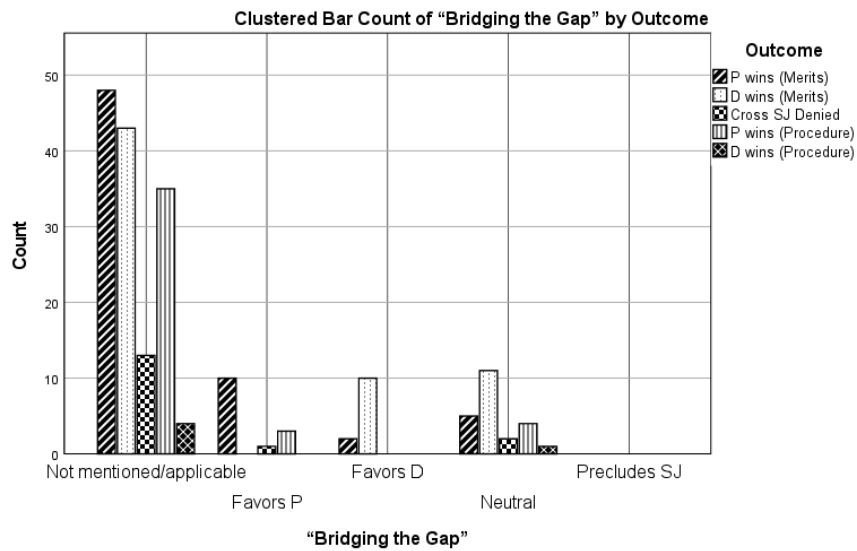
307. *DJ Direct, Inc. v. Margaliot*, 512 F. Supp. 3d 396, 411 (E.D.N.Y. 2021).

308. *Fortune Dynamic, Inc. v. Victoria’s Secret Stores Brand Mgmt., Inc.*, 618 F.3d 1025, 1038 (9th Cir. 2010).

309. *Brookfield Commc’ns, Inc. v. W. Coast Ent. Corp.*, 174 F.3d 1036, 1060 (9th Cir. 1999).

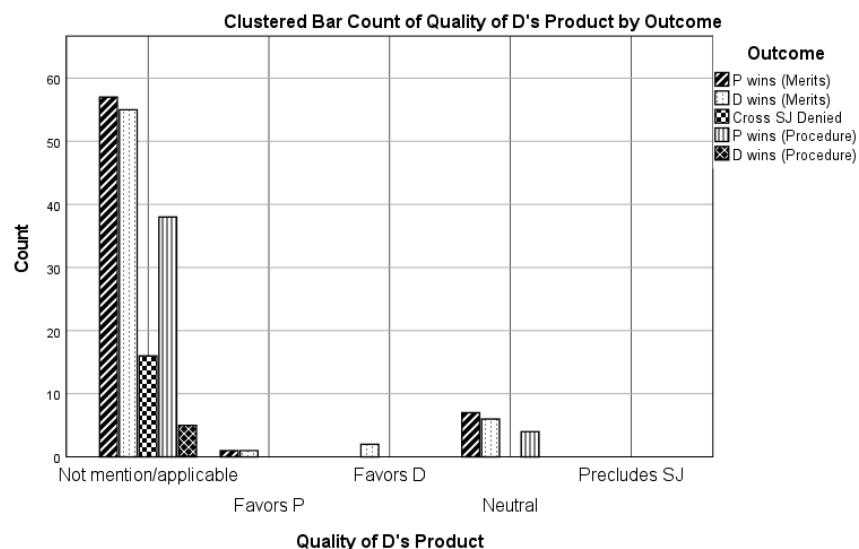
310. *Id.*

311. *Id.*

Figure 12: "Bridging the Gap" by Outcome

Similarly, the quality of the defendant's goods is closely related to competitive proximity. The quality factor featured even more infrequently, and 11% less than "bridging the gap." In 1% of all cases, the courts favored plaintiffs on the quality factor. When they did, plaintiffs won 50% of the time. In 1% of all cases, the courts favored defendants on the "bridging the gap" factor. When the court favored defendants, they won 100% of the time. Of course, these figures should be seen in the context of the very small sample size.

Figure 13: Quality of D's Product by Outcome



Courts look to the product, the relevant market, and potential consumers.³¹² Product proximity overlaps substantially with marketing and advertising channels and should be subsumed within those channels. For this reason, product proximity can serve as an omnibus factor for other factors such as the relative quality of goods sold, “bridging the gap” from the perspective of the relevant public (rather than from the legitimate aspirations of the trademark owner), and similarity of distribution channels.

One court in the dataset was exemplary in defining the relevant consumer market.³¹³ To determine whether that market included potential commercial and government customers, it examined the trademark owner’s revenue sources, proposals it sent to two potential customers, and the defendant’s exploratory acquisition of the trademark owner to conclude that both parties targeted similar customers.³¹⁴ Despite this fact-intensive inquiry, the court notably

312. *Best Cellars, Inc. v. Grape Finds at Dupont, Inc.*, 90 F. Supp. 2d 431, 456 (S.D.N.Y. 2000).

313. *Ironhawk Techs., Inc. v. Dropbox, Inc.*, 994 F.3d 1107, 1117 (9th Cir. 2021) (“Before addressing the *Sleekcraft* factors, we must define the relevant consumer market because ‘a court conducting a trademark analysis should focus its attention on the relevant consuming public.’” (quoting *Rearden LLC v. Rearden Com., Inc.*, 683 F.3d 1190, 1214 (9th Cir. 2012))).

314. *Id.* at 1117–18.

concluded that “a reasonable jury could find that [plaintiff’s] potential consumers include commercial customers.”³¹⁵

Market definition was an interesting issue that arose in a few cases in the dataset. Though similar, courts distinguish between market definition under trademark and antitrust law. For example, one case in the dataset reported a plaintiff asserting that “courts have looked to antitrust law . . . to find goods competitive where they are ‘either identical or available substitutes for each other.’”³¹⁶ Disagreeing with this assessment, the court responded that “the question . . . is not whether [the defendant’s] conduct impair[ed] competition in the marketplace[,] but whether it . . . infringed” upon a protected interest in the plaintiff’s trademark.³¹⁷

D. Summing It Up

In sum, the eight *Polaroid* Factors can be efficiently subsumed into a troika of actual confusion, mark similarity, and competitive proximity. The table below shows the troika being the most prominent factors. They also deliver consistent win rates to plaintiffs if the particular mark favors them, at between 45% to 54%, mapping almost exactly to Priest-Klein’s 50% figure discussed in Section I.A.

315. *Id.* at 1118.

316. *Flower Mfg., LLC v. CareCo, LLC*, 466 F. Supp. 3d 797, 812 (N.D. Ohio 2020) (emphasis omitted).

317. *Id.* (“It is fair to say that trademark laws were enacted for the protection of the competitor who owns a mark and not for protection of competition in the marketplace in general.”).

Figure 14: Revised Likelihood of Confusion Factors

<i>Polaroid</i> Factors	Rules of Thumb	Frequency (%)	Plaintiff's Win Rate (%)
Strength of the plaintiff's mark	Discarded	70	46
Similarity of plaintiff's and defendant's marks	Retained	85	47
Competitive proximity of products or services	Retained	73	45
Likelihood that plaintiff will "bridge the gap" and offer a product like a defendant's	Covered by competitive proximity of products or services	25	71
Actual confusion	Retained	74	54
Defendant's good faith	Discarded	66	52
Quality of defendant's product	Covered by competitive proximity of products or services	11	50
Buyer sophistication	Discarded	46	49

More importantly, the troika moves trademark doctrine a step in the right direction by limiting ad hoc fact-finding. However, the troika alone is incomplete. Mark McKenna and Mark Lemley warn that unless we can "identify more specifically the types of relationships that could give rise to actionable confusion, there is no logical stopping point for trademark protection."³¹⁸ The converse is also true—we also need to identify safe harbors. It is difficult even for savvy parties to predict the outcome in advance and resolve disputes early in proceedings, placing swathes of activity at significant risk.³¹⁹

318. See Mark A. Lemley & Mark McKenna, *Irrelevant Confusion*, 62 STAN. L. REV. 413, 439 (2010).

319. David S. Welkowitz, *The Virtues and Vices of Clarity in Trademark Law*, 81 TENN. L. REV. 145, 148 (2013) ("Because the level and even the existence of confusion is difficult to predict in advance, partly due to the uncertainties built into trademark law's test for confusion, those who would engage in valued activity must do so at significant risk.").

Simplifying confusion benefits other aspects of trademark law. For example, trademark law's first sale doctrine also permits some marketplace confusion by letting others sell used or reconditioned goods bearing the mark.³²⁰ Nominative fair use may likewise fold the likelihood of confusion standard into its analysis.³²¹ What is "fair" implicates the confusion arising from using the offending mark—whether the defendant only used as much as necessary of the plaintiff's mark—which in turn impacts the vagueness of the likelihood of confusion standard.³²² The same issue arises with expressive trademark uses³²³ and the legality of keyword advertising.³²⁴ Fair use is the focus of the next Section.

E. A Word on Fair Use

As trademarks expand beyond source identification, they seed public discourse with their communicative value.³²⁵ Trademark owners obtain rights with inchoate boundaries. When the public interacts with a trademark, the mark may imbue with collective meaning. This collective meaning has social value, and in appropriate instances, the law should offer them categorical protection from lawsuits.³²⁶

320. See *Nitro Leisure Prods., L.L.C. v. Acushnet Co.*, 341 F.3d 1356, 1362–64 (Fed. Cir. 2003) (consumer confusion as benchmark for applying the first sale doctrine).

321. *Toyota Motor Sales, U.S.A., Inc. v. Tabari*, 610 F.3d 1171, 1175–76 (9th Cir. 2010) (asking whether (1) the product was readily identifiable without use of the mark; (2) defendant used more of the mark than necessary; or (3) defendant falsely suggested he was sponsored or endorsed by the trademark holder).

322. *E.g.*, *KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc.*, 543 U.S. 111, 123 (2004) (confusion relevant to whether descriptive use is "fair").

323. See *Rogers v. Grimaldi*, 875 F.2d 994, 999 (2d Cir. 1989) (adopting balancing test that asks whether the use of a trademark as the title of an expressive work is artistically relevant to the underlying work and, if so, whether "the title explicitly misleads as to the source or the content of the work").

324. See *Network Automation, Inc. v. Advanced Sys. Concepts, Inc.*, 638 F.3d 1137, 1154 (9th Cir. 2011) (noting that keyword advertisements could be "confusingly labeled or not labeled at all" making how advertisements appear on the results page must be considered).

325. See Alex Kozinski, *Trademarks Unplugged*, 68 N.Y.U. L. REV. 960, 973–74 (1993) (noting how businesses inject the "effervescent qualities" of trademarks "into the stream of communication with the pressure of a firehose by means of mass media campaigns").

326. See, *e.g.*, William McGeveran & Mark P. McKenna, *Confusion Isn't Everything*, 89 NOTRE DAME L. REV. 253, 301–06 (2013) (proposing categorical exclusions for some favored uses).

Communication relies on a plethora of legally protected words, graphics, sounds, and smells.³²⁷ Beyond computers or smartphones, APPLE may represent a nonconformist hip lifestyle compared with users of LENOVO's more staid business offerings. Trademarks become tools of communication and expression, and the public helps shape their boundaries as they become symbols that embody culture itself.³²⁸

Trademark owners may be anxious to protect themselves from uses that dilute the value of a household logo or name even when consumers are not confused. Between 2019 and 2021, Apple filed 215 trademark oppositions, targeting small companies and nonprofits that have nothing to do with providing technology products or services, including an Indian food blog and a public school.³²⁹ In these mass-produced, boilerplate-worded oppositions, Apple has argued that "Apple marks are so famous and instantly recognizable" that other trademarks will weaken the strength of its brand or cause the "ordinary consumer to believe that applicant is related to, affiliated with or endorsed by Apple."³³⁰ While Apple protested that this is simply what the law "requires," Professor Christine Farley has called them "bullying tactics."³³¹ Whatever the case, the impact is clear; Apple has been successful in preventing registration of a wide variety of marks. When faced with an opposition by Apple, applicants expressly withdrew their applications 17% of the time and failed to respond and subsequently defaulted 59% of the time.³³² The Tech Transparency Report noted only one win against Apple—by the U.S. government:

327. Diamond & Franklyn, *supra* note 165, at 2031.

328. Barton Beebe, *The Semiotic Analysis of Trademark Law*, 51 UCLA L. REV. 621, 624 (2004) (arguing trademark law is both an economic doctrine and "a semiotic doctrine elaborating the principles of sign systems, of language").

329. See *Apple's Trademark 'Bullying' Targets Small Businesses, Nonprofits* TECH TRANSPARENCY PROJECT (Mar. 11, 2022) [hereinafter TECH TRANSPARENCY PROJECT], <https://www.techtransparencyproject.org/articles/apples-trademark-bullying-targets-small-businesses-nonprofits> [<https://perma.cc/96QR-UZBT>] ("Many choose to simply give up rather than take on a mega-corporation with a market value of \$2.5 trillion, but in doing so, they lose whatever funds they invested in designing their logo and hiring a lawyer to deal with the trademark application.").

330. Ryan Mac & Kellen Browning, *Apps and Oranges: Behind Apple's 'Bullying' on Trademarks*, N.Y. TIMES (Mar. 11, 2022) <https://www.nytimes.com/2022/03/11/technology/apple-trademarks.html> [<https://perma.cc/LEV8-RPEE>].

331. *Id.* For an interesting discussion on how to prevent "bullying" in trademark law, see Stacey Dogan, *Bullying and Opportunism in Trademark and Right-of-Publicity Law*, 96 B.U. L. REV. 1293 (2016) (examining various instances of trademark bullying, including by Monster Energy, one of the most infamous trademark bullies).

332. See TECH TRANSPARENCY PROJECT, *supra* note 329 ("Of the 118 North American cases analyzed by TTP, 76 have been decided in Apple's favor, with a complete defeat

Apple appears to have been stopped in only one instance: when it challenged the U.S. government. The Department of Energy registered an online research service called Pages in 2020. Apple owns the trademark Pages for its word processor and opposed the agency's trademark on those grounds. The Department's response was brief, and its main defense took the form of a one-sentence declaration: "There is no likelihood of confusion, mistake, or deception between [Apple's] marks and DOE's PAGES mark." After more than a year of negotiations, Apple agreed to withdraw its objections, without DOE making any change to its application. In TTP's dataset, no other applicant was able to completely beat back Apple.³³³

While the costs of invading free speech and other interests are high, the costs of being overly permissive in expressive use cases cause only minimal harm.³³⁴ Research on brand extensions shows owners are rarely "harmed by consumers' mistaken association of unrelated products."³³⁵ Consumers rarely alter how they view the brand quality when they encounter criticism about other products offered under that mark.³³⁶ The negative impact stays with the related products but does not corrupt a positive view of the owner's line of products.³³⁷

For this reason, Mark McKenna has warned against reflexively prohibiting every form of confusion.³³⁸ Instead, trademark law should only be concerned with confusion that influences consumer decision-making. Bone also cautions that economic concerns over confusion should be distinguished from penalizing intentional deception without evidence of consumer confusion.³³⁹ Similarly, Lisa Ramsey flags the need to safeguard free speech interests in the face of

of the other parties' proposed logos or trademarks. None of these cases saw a full trial before TTAB but were withdrawn or abandoned by the applicants amid Apple's pressure.").

333. *Id.*

334. William McGeeveran, *The Trademark Fair Use Reform Act*, 90 B.U. L. REV. 2267, 2286 (2010).

335. Lemley & McKenna, *supra* note 318, at 429.

336. *See id.* at 429.

337. *See id.* at 430 ("Consumers, in other words, are smart enough to distinguish different products and hold different impressions of them.").

338. Mark P. McKenna, *A Consumer Decision-Making Theory of Trademark Law*, 98 VA. L. REV. 67, 73 (2012).

339. Robert G. Bone, *Taking the Confusion out of "Likelihood of Confusion": Toward a More Sensible Approach to Trademark Infringement*, 106 NW. L. REV. 1307, 1377 (2012).

encroaching trademark enforcement.³⁴⁰ One way to do this is by fortifying fair use.

Fair use is currently regarded as an affirmative defense.³⁴¹ That generally precludes the pretrial disposition of the case.³⁴² In the interest of early off-ramping cases, it is perhaps fortunate then that fair use rarely arose in the dataset (6%), with equal probability that a court would eventually find in favor of either plaintiff (42%) or defendant (50%). It also provides empirical evidence that converting fair use from an affirmative defense to a safe harbor would create a powerful tool to fend off trademark trolls without appreciably disrupting day-to-day trademark practice.

Safe harbors offer advantages over attempts to prescribe clear rules. These include improving predictability and ease of determination, allowing courts to resolve issues sooner in the litigation process. Here, Gideon Parchomovsky and Alex Stein make a more general point that “[r]eplacing these criteria with rules that will lay down irrebuttable presumptions of consumer confusion, or lack thereof, could make litigation over trademarks cheaper than it presently is.”³⁴³ The case is over as soon as the defendants demonstrate a basic fact.³⁴⁴

Safe harbors exist within trademark law, specifically the likelihood of confusion tests. For instance, the law does not protect functional product designs to avoid giving plaintiffs an advantage against rivals unrelated to the plaintiff’s reputation.³⁴⁵ Similarly, the law keeps plaintiffs on a leash to not monopolize trademarks with descriptive words and receive protection for generic terms.³⁴⁶ Expressive uses for

340. See Lisa P. Ramsey, *Increasing First Amendment Scrutiny of Trademark Law*, 61 SMU L. REV. 381, 384–85 (2008) (advocating rigorous First Amendment rights of trademark laws to protect free speech).

341. See, e.g., *KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc.*, 543 U.S. 111, 123 (2004) (finding that confusion is relevant to whether descriptive use is “fair”); *New Kids on the Block v. News Am. Publ’g, Inc.*, 971 F.2d 302, 308 (9th Cir. 1992) (noting that confusion is relevant to nominative fair use).

342. *Kelly-Brown v. Winfrey*, 717 F.3d 295, 308 (2d Cir. 2013) (“fair use . . . requires consideration of facts outside of the complaint and thus is inappropriate to resolve on a motion to dismiss.”); see also 2 MCCARTHY, *supra* note 2, § 11:49 (“Because classic fair use is an affirmative defense, it is normally not appropriate for consideration on a . . . motion to dismiss for failure to state a claim.”).

343. Gideon Parchomovsky & Alex Stein, *Catalogs*, 115 COLUM. L. REV. 165, 178 (2015).

344. See Welkowitz, *supra* note 319, at 168 (referencing Fed. R. Evid. 301).

345. 15 U.S.C. § 1052(e)(5).

346. See *Abercrombie & Fitch Co. v. Hunting World, Inc.*, 537 F.2d 4, 9–10 (2d Cir. 1976) (explaining the limitations on generic and descriptive marks).

commentary, parody, or education should fall within safe harbors.³⁴⁷ Critiquing products or corporate behavior requires us to use them.³⁴⁸ Therefore, the first safe harbor should be expressive uses of protected trademarks.

The second safe harbor is referential uses of trademarks. Nominative fair use (referring to the trademark holder or its products) should not trigger liability.³⁴⁹ For example, rivals and repair services need to make referential uses to compete and advertise their services to the public.³⁵⁰ The law currently recognizes comparative use as a defense, but it should go further and offer a safe harbor to these uses.³⁵¹

Recognizing that the *Polaroid* factors are a “bad fit” in nominative fair use cases, one court instructed that future courts should consider fair use alongside the *Polaroid* factors when considering a claim of nominative fair use.³⁵² Likelihood of confusion is relevant to determining whether the use is objectively fair and whether defendants use the term “other[] than as a mark.”³⁵³ Likewise, nominative fair use folds confusion into determining whether an expressive use “explicitly misleads” consumers or whether the use falsely suggests a source or sponsorship.³⁵⁴

347. See Andy Greene, *Nathan Fielder Talks ‘Dumb Starbucks’ and Pranking Instagram*, ROLLING STONE (July 24, 2014), <http://www.rollingstone.com/movies/news/nathan-fielder-talks-dumb-starbucks-and-pranking-instagram-20140724> [<https://perma.cc/Q3BL-R96R>] (noting that parody laws allowed comedian Nath Fielder to “open up a near perfect replicate of a Starbucks just as long as he put the word ‘dumb’ before everything in the store, down to CDs labeled ‘Dumb Nora Jones Duets’”).

348. See *New Kids on the Block v. News Am. Publ’g, Inc.*, 971 F.2d 302, 307 (9th Cir. 1992) (“Much useful social and commercial discourse would be all but impossible if speakers were under threat of an infringement lawsuit every time they made reference to a person, company or product by using its trademark.”).

349. See *e.g.*, *Zatarains, Inc. v. Oak Grove Smokehouse, Inc.*, 698 F.2d 786, 788 (5th Cir. 1983) (rivals allowed to use “fish fry” to describe their own batter mixes even when doing so creates some likelihood of confusion with owners’ FISH-FRI trademark).

350. See *e.g.*, *Toyota Motor Sales, U.S.A., Inc. v. Tabari*, 610 F.3d 1171, 1180–82 (9th Cir. 2010) (allowing automobile broker specializing in facilitating Lexus purchases to use LEXUS mark as part of domain name).

351. See *e.g.*, *Smith v. Chanel, Inc.*, 402 F.2d 562, 563 (9th Cir. 1968) (holding that truthful comparative advertising is not trademark infringement).

352. *Int’l Info. Sys. Sec. Certification Consortium, Inc. v. Sec. Univ., LLC*, 823 F.3d 153, 168 (2d Cir. 2016); see also *Coty Inc. v. Excell Brands, LLC*, 277 F. Supp. 3d 425, 457 (S.D.N.Y. 2017) (“Acourt considering a claim of nominative fair [use] should consider three factors in addition to the standard *Polaroid* factors.”).

353. See *e.g.*, *KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc.*, 543 U.S. 111, 118 (2004).

354. William McGeveran, *Rethinking Trademark Fair Use*, 94 IOWA L. REV. 49, 95–97, 100–04 (2008).

Safe harbors like those for expressive and descriptive uses allow courts to dispose of the likelihood of confusion cases more simply and justly. For example, uses that mirror the conventional way descriptive terms are used in ordinary language give prospective users an advantage in establishing the protected use and exiting litigation early, thereby avoiding high litigation costs. In addition, they help carve out pockets of strong protection and guide the development of trademark rights in other areas such as merchandising rights, without giving owners the right to rely upon the likelihood of confusion to justify its approval. Within this framework, it is also worth considering a safe harbor beyond descriptive or expressive fair uses that provide small businesses and nonprofits like those described above with an effective and low-cost way to deflect policing by overzealous trademark owners.

III. OBSERVATIONS AND IMPLICATIONS

The final Part addresses three issues. First, it observes that courts combine factors and analyze them together. Occasionally they do so overtly. Factor folding occurs across all the factors, including the troika of mark similarity, actual confusion, and competitive proximity. A strong showing on one factor may prevent the need to show another. Second, decision-makers tend to start limiting the factors that they choose to consider when confronted with complex decision processes. At some point, decision-makers will stop analyzing new information and instead commit to a decision and then work backward to vindicate it. This adaptation has allowed 63% of litigants to receive an early resolution on the merits. Third, it explains how the empirical analysis provides a blueprint for algorithmic adjudication using AI, taking the reader from conception to execution to identifying and addressing its limitations.

A. Factor Folding

While likelihood of confusion factors may present themselves as discrete categories, the dataset reveals that courts do not regard them as such. Courts instead combine factors and analyze them together. This is called “folding” and in likelihood of confusion analyses the courts notoriously fold the factors together, using the presence of one factor as a proxy analysis for another.

For instance, in *J-B Weld Co., LLC v. Gorilla Glue Co.*,³⁵⁵ the Eleventh Circuit used similarity as a proxy for intent.³⁵⁶ Instead of making an adverse finding outright, a court may sometimes shift the plaintiff's burden of proof to require the defendant instead to disprove bad faith.³⁵⁷ The direction of a court's substitution bias is not a one-way street. On occasion, courts leaned on lack of evidence of actual confusion to vindicate imitation of successful product features.³⁵⁸ In this way, the likelihood of confusion factors operate not as independent elements along orthogonal lines but as a sliding scale: the more closely the products compete, the more likely it is that a new product whose design arrogates the atypical qualities of the old product will confuse consumers.

Sometimes the combination is obvious. For instance, courts treat actual confusion as an indicator of mark strength.³⁵⁹ One court explained that "[i]f buyers are confused between two sources, then this also means that they must have recognized plaintiff's designation as a trademark and associated it only with the plaintiff."³⁶⁰ Another court observed that "where the parties' marks are identical and their goods are in very close competitive proximity, a highly sophisticated consumer may be the most vulnerable to confusion."³⁶¹

This blending was not confined to the likelihood of confusion factors but extended to fair use. One court explained actual confusion gets to "the heart of the nominative fair use situation."³⁶² At other times, the logical connection is more tenuous, suggesting a negative

355. 978 F.3d 778, 790 (11th Cir. 2020).

356. *Id.* ("[W]here a defendant attempts to copy a plaintiff's product 'as closely as possible' and uses the plaintiff's product design as a model, it may be 'inferred that [defendant] purposely chose a mark which was very similar to [plaintiff's] in order to benefit from the reputation [plaintiff]'s mark had already achieved.'").

357. *Kiki Undies Corp. v. Promenade Hosiery Mills, Inc.*, 411 F.2d 1097, 1101 (2d Cir. 1969) ("[W]here the allegedly infringing mark is identical to the registered mark, and its use began subsequent to the plaintiff's trade-mark registration, the defendant must carry the burden of explanation and persuasion.").

358. *George Basch Co. v. Blue Coral, Inc.*, 968 F.2d 1532, 1541 (2d Cir. 1992) ("Absent confusion, imitation of certain successful features in another's product is not unlawful . . .").

359. *Vital Pharms., Inc. v. Monster Energy Co.*, No. 19-60809-CIV, 2021 WL 3371942, at *43 (S.D. Fla. Aug. 3, 2021) (stating that several courts have held that "actual confusion is an indicium of secondary meaning"); *Am. Sci. Chem., Inc. v. Am. Hosp. Supply*, 690 F.2d 791, 793 (9th Cir. 1982).

360. 2 MCCARTHY ON TRADEMARKS § 15:11; *see also* *Popular Bank of Fla. v. Banco Popular de P.R.*, 9 F. Supp. 2d 1347, 1358 (S.D. Fla. 1998) (making the same point).

361. *Kohler Co. v. Bold Int'l FZCO*, 422 F. Supp. 3d 681, 730 n.20 (E.D.N.Y. 2018).

362. *Juul Labs, Inc. v. 4X PODS*, 509 F. Supp. 3d 52, 64 (D.N.J. 2020).

form of coherence-based reasoning. For instance, in ascertaining mark similarity, one court considered the nature of the purchasing process, reasoning that where marks are similar but used in different contexts or on different visual displays, the risk of confusion is minimized.³⁶³

This practice of factor folding happens across all factors, including with the trio of key factors of mark similarity, actual confusion, and competitive proximity, where a strong showing on one factor may be sufficient. A strong showing on one factor may prevent the need to show another.³⁶⁴ Factors thus trump each other, with competitive proximity often trumping mark similarity without explaining why one factor should take precedence over another.³⁶⁵

Notably, only 9% of cases in the dataset expressly acknowledge “folding” factors. Most do not, regardless of the procedural posture in the case. This makes it more difficult for appellate courts and commentators to hold lower courts accountable for their analysis when this “folding” occurs. This phenomenon underscores the importance of minimizing coherence-based reasoning by having courts focus on a few factors when making likelihood of confusion determinations.

B. Early Off-Ramps

Courts generally agree that “application of the factors is a highly fact-intensive inquiry both as to the assessment of the evidence concerning each factor and as to the overall synthesis of factors and the evidence.”³⁶⁶ A context-specific inquiry guides courts towards the material aspects of product source or affiliation germane to the consuming public’s understanding.³⁶⁷ Given their marching orders, one might expect judges to weigh the likelihood of confusion factors

363. *Sensient Techs. Corp. v. SensoryEffects Flavor Co.*, 613 F.3d 754, 765 (8th Cir. 2010); *E.A. Sween Co. v. A & M Deli Express, Inc.*, No. 17 CV 2514, 2018 WL 1283682, at *4 (E.D.N.Y. Mar. 9, 2018).

364. *See e.g., ConAgra, Inc. v. George A. Hormel & Co.*, 990 F.2d 368, 371 (8th Cir. 1993) (“[W]hen ‘products are closely related, less similarity in trademarks is necessary to support a finding of infringement.’” (quoting *SquirtCo v. Seven-Up Co.*, 628 F.2d 1086, 1091 (8th Cir. 1980))); *Select Comfort Corp. v. Baxter*, 996 F.3d 925, 934 (8th Cir. 2021) (“[T]he relative importance of any given factor is influenced greatly by how the other factors might apply.”).

365. *Reply All Corp. v. Gimlet Media, LLC*, 843 F. App’x 392, 396 (2d Cir. 2021) (“[W]hile the two marks undoubtedly share aural and typographic similarities, they are unlikely to appear in the marketplace in a similar manner.”).

366. *Select Comfort Corp.*, 996 F.3d at 933–34.

367. *Id.* at 934 (“Common sense is inherent in the factors, and the factors, properly applied, should try to capture a holistic view of the normal experiences for any given industry, product, or service.”).

carefully.³⁶⁸ However, as seen in this study, that is not what happens in practice, nor indeed more generally.³⁶⁹

When confronting complex decision processes, decision-makers tend to limit the factors they consider.³⁷⁰ After a certain point, judges will stop analyzing new information, instead committing to their decision first and then working backwards to rationalize it. Some courts opt for a holistic weighing of the factors rather than attempting piecemeal arithmetic.³⁷¹ Others emphasize case-by-case determination, and in so doing, underscore flexibility in applying a multitude of factors.³⁷²

To resist a movant's summary judgment motion, the non-moving party must establish, through pleadings, depositions, answers to interrogatories, admissions, and affidavits in the record that there is no genuine issue of material fact.³⁷³ To put it differently, summary judgment needs to be based on undisputed material facts that show there is "only one conclusion a trier of fact could reasonably draw."³⁷⁴ The "factors require a fact-intensive inquiry not suitable for a motion to dismiss."³⁷⁵ Some appellate courts caution district courts to only grant summary judgment on the issue of likelihood of confusion "sparingly."³⁷⁶

However, expedient determinations serve the ends of justice for both sides in litigation. Summary judgments provide a quick and

368. See *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309, 322 (S.D.N.Y. 2003) (noting the tendency towards this type of application).

369. Anthony E. Chavez, *Using Legal Principles to Guide Geoengineering Deployment*, 24 N.Y.U. ENVTL. L.J. 59, 93 (2016) ("Decision makers, however, often do not apply multi-factor—or multi-principle—tests as they are intended.").

370. Beebe, *supra* note 6, at 1601.

371. *Noble v. United States*, 231 F.3d 352, 359 (7th Cir. 2000).

372. John S. Applegate, *Worst Things First: Risk, Information, and Regulatory Structure in Toxic Substances Control*, 9 YALE J. REG. 277, 302 (1992).

373. *ServPro Intell. Prop., Inc. v. Blanton*, 451 F. Supp. 3d 710, 722 (W.D. Ky. 2020).

374. *Health Net v. U.S.A. Healthnet, Inc.*, No. CV 92-3925 KN, 1993 WL 209558, at *1 (C.D. Cal. May 12, 1993).

375. *GeigTech E. Bay LLC v. Lutron Elecs. Co.*, 352 F. Supp. 3d 265, 285 (S.D.N.Y. 2018); see also *Eliya, Inc. v. Kohl's Dep't Stores*, No. 06 Civ 195, 2006 WL 2645196, at *13 n.2 (S.D.N.Y. Sept. 13, 2006) ("[A]n application of the so-called *Polaroid* factors on this motion to dismiss would be inappropriate because it would involve premature factfinding.").

376. *Fortune Dynamic, Inc. v. Victoria's Secret Stores Brand Mgmt., Inc.*, 618 F.3d 1025, 1039 (9th Cir. 2010); see also *Rearden LLC v. Rearden Com., Inc.*, 683 F.3d 1190, 1210 (9th Cir. 2012) ("Given the open-ended nature of this multi-prong inquiry, it is not surprising that summary judgment on 'likelihood of confusion' grounds is generally disfavored.").

inexpensive off-ramp for parties to dispose of a case when no real issues call for a trial. The ability of courts to wield this important judicial tool protects defendants against frivolous lawsuits and plaintiffs from incurring unnecessary costs.³⁷⁷ Streamlining the test by consolidating and trimming down the factors will enable courts to get to the heart of the inquiry expeditiously. Simplifying the likelihood of confusion lowers the temperature and makes it easier for owners to determine when to protect their interests.

Many courts are more willing to move ahead with summary judgments even where factors are in dispute and the evidence is not obvious on the basis that “as with any other issue of fact, summary judgment remains appropriate when no jury reasonably could have ruled in the non-moving party’s favor.”³⁷⁸ Others are willing to do so when most of the relevant factors weigh in the movant’s favor, including at least one “key factor.”³⁷⁹ Yet others maintain “a finding of a likelihood of confusion ‘need not be supported by a majority’ of the digits.”³⁸⁰ Non-movants resisting summary judgment must show “how additional discovery on these issues would create a genuine issue of fact” material to movants’ claim for trademark infringement.³⁸¹

In a Ninth Circuit case from the dataset, the court held as a matter of law that the trademark owner was entitled to summary judgment where the marks were identical, the goods were related, and the marketing channels overlapped.³⁸² A small set of key factors helps structure the likelihood of confusion inquiry and gives notice of pertinent issues and relevant evidence; this creates a more solid basis

377. Elaine Kussurelis, *Canada’s Summary Trial Procedure: A Viable Alternative to Summary Judgment on Trademark Likelihood of Confusion Actions in the United States*, 50 U. MIAMI INTER-AM. L. REV. 165, 168 (2019) (observing summary judgments “can be a powerful trademark litigation weapon for either plaintiffs or defendants”).

378. *RXD Media, LLC v. IP Application Dev. LLC*, 986 F.3d 361, 375 (4th Cir. 2021); see also *EST Inc. v. Royal-Grow Prods., LLC*, 526 F. Supp. 3d 943, 956–57 (D. Kan. 2021) (“The evidence is far from one-sided and leads to no obvious answer.”); *Collins v. U.S. Dep’t of Veterans Affs.*, 497 F. Supp. 3d 885, 894 (S.D. Cal. 2020) (“[S]ummary judgment is still proper in trademark infringement cases where, as here, no genuine issue of material fact exists.”).

379. See, e.g., *RXD Media, LLC*, 986 F.3d at 375 (“Based on the record before us, we hold that a jury could not have reasonably concluded that RXD’s use of the ‘ipad’ mark was unlikely to cause consumer confusion.”).

380. *Future Proof Brands, LLC v. Molson Coors Beverage Co.*, 982 F.3d 280, 298 (5th Cir. 2020).

381. *Better Angels Soc’y, Inc. v. Inst. for Am. Values, Inc.*, 419 F. Supp. 3d 765, 780 (S.D.N.Y. 2019).

382. *Nissan Motor Co. v. Nissan Comput. Corp.*, 378 F.3d 1002, 1019 (9th Cir. 2004).

for predicting case outcomes and may even be sufficient to justify an inference that there is a likelihood of confusion.³⁸³

Similarly, courts can also rule on a motion to dismiss under Rule 12(b)(6) if the complaint contains facts that state a claim to relief that is plausible on its face that the plaintiff's claims are barred as a matter of law.³⁸⁴ Where plaintiffs can substantiate the plausibility of their claims, courts will deny the motion to dismiss.³⁸⁵ Plaintiffs have successfully done so on competitive proximity,³⁸⁶ actual confusion,³⁸⁷ mark strength,³⁸⁸ mark similarity, and "bridging the gap."³⁸⁹

This Article reveals that courts rely on a small number of factors to economize their decisions to give parties an early off-ramp. With either party as the movant or cross-motion, summary judgments comprised 48% of cases in the dataset. Motions to dismiss by either party made up 10% of cases, and other postures, mostly preliminary injunction motions, made up 21%. These collectively indicate that 79% of litigants seek an early resolution. Of these, 63% received a resolution with either plaintiff or defendant winning on the merits.

Some district courts treat the likelihood of confusion as a question of fact, requiring proof of each element of each factor and categorically precluding summary judgment.³⁹⁰ Others treat the

383. *Future Proof Brands, LLC*, 982 F.3d at 298.

384. *Uber Inc. v. Uber Techs., Inc.*, 521 F. Supp. 3d 455, 462 (S.D.N.Y. 2021); *see also id.* at 464 ("At the Rule 12(b)(6) stage, the *Polaroid* analysis is limited to the facts alleged in the Complaint and any documents integral thereto.").

385. *See Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009) (holding that "[t]o survive a motion to dismiss," a complaint need only "contain sufficient matter, accepted as true, to 'state a claim to relief that is plausible on its face.'" (quoting *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 570 (2007))).

386. *Uber Inc.*, 521 F. Supp. 3d at 464 ("It suffices to note that the Complaint plausibly alleges that defendants' services are in competitive proximity with the plaintiff's graphic design and marketing services.").

387. *Id.* ("[T]he Complaint's descriptions of confusion among businesses, official bodies and members of the public provide some factual support for the plausibility of plaintiff's claims.").

388. *Id.* ("Among the other *Polaroid* factors, the Complaint plausibly alleges that plaintiff's Uber mark is arbitrary and distinctive, that the parties' marks are similar, and that defendants' putative entry into the display-advertising market may result in plaintiff bridging the 'gap' between its services and those of defendants.").

389. *Id.*

390. *Flower Mfg., LLC v. CareCo, LLC*, 466 F. Supp. 3d 797, 820 (N.D. Ohio 2020) ("Because each of Flower's claims requires proof of that element, they cannot survive summary judgment.").

likelihood of confusion as a matter of law, whether outright³⁹¹ or paying lip service to its factual dimensions.³⁹² Appellate courts are split along the same lines, either reviewing the lower courts' "ultimate conclusion about likelihood of success for clear error,"³⁹³ or "de novo, using the same legal standards [the lower court] employed."³⁹⁴

Even among those who profess fidelity to the fact/law distinction in theory, applying that distinction in practice is not easy. As the Sixth Circuit put it, "[a]ny dispute about the evidence that pertains to the eight factors presents a factual issue . . . [and] whether a given set of foundational facts establishes a likelihood of confusion is a legal conclusion."³⁹⁵ In contrast, the Second Circuit expressed that determining whether one of the *Polaroid* factors favors one party or another is a legal judgment reviewed de novo.³⁹⁶

The data also revealed a steady increase in the affirmance rate of lower courts' decisions between 2016 and 2021. This is because appellate courts generally defer to lower court finding of facts but give less deference to questions of law.³⁹⁷ But, more significantly, appellate courts seem either unaware or complicit in the practice of lower courts folding some factors and ignoring others. It would be interesting and worth further study to see whether this is a practice of "wink-and-nod" between the lower and appellate courts or if this state of affairs was purely coincidental.

391. *FCOA, LLC v. Foremost Title & Escrow Servs., LLC*, 416 F. Supp. 3d 1381, 1387 (S.D. Fla. 2019) ("[I]n trademark infringement cases, courts in this Circuit have decided the issue of likelihood of confusion as a matter of law.").

392. *Yellowfin Yachts, Inc. v. Barker Boatworks, LLC*, 898 F.3d 1279, 1289 (11th Cir. 2018) ("Although likelihood of confusion is a question of fact, it may be decided as a matter of law.").

393. *Future Proof Brands, LLC v. Molson Coors Beverage Co.*, 982 F.3d 280, 298 (5th Cir. 2020); *AWGI, LLC v. Atlas Trucking, Co., LLC*, 998 F.3d 258, 264 (6th Cir. 2021) ("We review the district court's finding of fact for clear error" (quoting *Premium Freight Mgmt., LLC v. PM Eng'g Sols., Inc.*, 906 F.3d 403, 406 (6th Cir. 2018))).

394. *Yellowfin Yachts, Inc.*, 898 F.3d at 1289.

395. *Progressive Distrib. Servs., Inc. v. United Parcel Serv., Inc.*, 856 F.3d 416, 427 (6th Cir. 2017); *see also* *Champions Golf Club, Inc. v. The Champions Golf Club, Inc.*, 78 F.3d 1111, 1116 (6th Cir. 1996) ("Whether there is a likelihood of confusion is a mixed question of fact and law.").

396. *Tiffany & Co. v. Costco Wholesale Corp.*, 971 F.3d 74, 86 (2d Cir. 2020) ("[I]nsofar as the determination of whether one of the *Polaroid* factors favors one party or another involves a legal judgment—which it often does—we must review that determination de novo.").

397. *Select Comfort Corp. v. Baxter*, 996 F.3d 925, 934 (8th Cir. 2021) ("[W]e review the likelihood of confusion determination as a finding of fact.").

Figure 15: Affirmance on Appeal

2016	2017	2018	2019	2020	2021
100.0%	87.5%	80.0%	84.6%	81.3%	81.8%

Scholars have long debated whether AI can replicate human legal reasoning.³⁹⁸ Edward Levi described how common law rules evolve in his classic text *An Introduction to Legal Reasoning*.³⁹⁹ Judges begin by identifying factors that have legal salience to explain case outcomes. Once those rules fail to yield sensible results, judges alter them accordingly.⁴⁰⁰ Legal reasoning rests on analogies, but this fact-specific method also makes developing coherence in the case of precedent an elusive task. For this reason, AI's ability to detect patterns in judicial opinions is of great interest to scholars.⁴⁰¹

At its heart, the likelihood of confusion inquiry seeks to ascertain the probability that a defendant's use of its trademarks will confuse consumers.⁴⁰² Making the likelihood of confusion more rule-like, both through the doctrinal reformation of the standard and through the application of AI, makes it easier for appeals courts to scrutinize and overturn deviant lower court decisions and allows lower courts to

398. See e.g., Stephen M. McJohn, *Review of Artificial Legal Intelligence*, 12 HARV. J.L. & TECH. 241, 243–44 (1998) (noting that the way neural networks learn through adjustment makes the technology capable of performing legal reasoning; however, the technology would not understand the reasons behind its conclusion, making it less valuable to the legal field).

399. Edward H. Levi, *An Introduction to Legal Reasoning*, 15 U. CHI. L. REV. 501, 501–03 (1948).

400. *Id.*; see also RONALD DWORKIN, *LAW'S EMPIRE* 400–03 (1986) (explaining the factors that a judge might weigh when considering whether to change the law prospectively, noting that they are bound by precedent and the integrity of law).

401. See, e.g., McJohn, *supra* note 398, at 241 (offering commentary on another scholar's contribution to discussions around the topic).

402. Leah Chan Grinvald, *Shaming Trademark Bullies*, 2011 WIS. L. REV. 625, 636 (2011) (“This liability standard refers to the probability (not the actuality or possibility) that consumers will be confused by the same or similar trademarks.”).

distinguish dubious precedent based on facts.⁴⁰³ Trademark's troika of actual confusion, mark similarity, and competitive proximity paves the road for AI to fill the final piece of the equation to simplify the likelihood of confusion.

C. Deploying Artificial Intelligence

Legal scholarship on AI and trademarks is scarce.⁴⁰⁴ It is surprising given AI's centrality in both the consumer marketing literature and trademark's centrality in IP protection.⁴⁰⁵ Sonia Katyal and Aniket Kesari argued "as a general matter, that AI should be of interest to anyone studying trademarks and the role that they play in economic decision making."⁴⁰⁶ They point to AI deployment by the government in trademark image recognition, classifying goods and services, and identifying descriptive terms.⁴⁰⁷ This Article explains how a dataset such as the one used here might be a rudimentary prototype for a grander form of AI-enabled likelihood of confusion analysis that courts and litigants might deploy in the future.

403. Frederick Schauer, *Formalism*, 97 YALE L.J. 509, 541–42 (1988) (noting errors are more easily detectable under rules).

404. Sonia K. Katyal & Aniket Kesari, *Trademark Search, Artificial Intelligence, and the Role of the Private Sector*, 35 BERKELEY TECH. L.J. 501, 504 (2020) ("[S]urprisingly, very little legal scholarship has addressed the potential role for AI in the context of trademarks."). The few examples available include Dev Gangjee, *Eye, Robot: Artificial Intelligence and Trade Mark Registers*, in TRANSITION AND COHERENCE IN INTELLECTUAL PROPERTY LAW (Niklas Bruun, Graeme B. Dinwoodie, Marianne Levin & Ansgar Ohly eds., 2021), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3467627; Anke Moerland & Conrado Freitas, *Artificial Intelligence and Trade Mark Assessment*, in ARTIFICIAL INTELLIGENCE & INTELLECTUAL PROPERTY (Jyh-An Lee, Reto M. Hilty, & Kung-Chung Liu eds., 2021), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3683807.

405. See *Trademarks, Copyright and Patents: Should Business Owners Really Care About IP?*, VARNUM (May 1, 2019), <https://www.varnumlaw.com/newsroom-publications-trademarks-copyrights-and-patents-why-business-owners-should-care-about-ip> [<https://perma.cc/VST8-XF56>] ("A trademark is one of the most important business assets that a company will ever own because it identifies and distinguishes the company and its products/services in the marketplace from its competitors.").

406. Katyal & Kesari, *supra* note 404, at 505 ("AI will fundamentally transform the trademark ecosystem, and the law will need to evolve as a result.").

407. See, e.g., DAVID FREEMAN ENGSTROM ET AL., GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 49–50 (2020), <https://www-cdn.law.stanford.edu/wp-content/uploads/2020/02/ACUS-AI-Report.pdf> [<https://perma.cc/SR9J-UEUN>] (describing a USPTO prototyping using deep learning model using an unsupervised approach to generate visually similar images from a database).

1. *Conception*

AI gives courts the capability to scour case reports to assess how past courts weighed effects and stress-test theories of confusion against real-world data.⁴⁰⁸ AI can match the results against depositions and other preprocessed evidence to provide quicker and more consistent analyses, unlike the binarily coded factors in this study.⁴⁰⁹ Principal component analysis can identify factors carrying the greatest weight in functions and zero in on the most important dimensions of datasets to show the stampeding likelihood of confusion factors.⁴¹⁰

AI expands the scope of cases so that courts can dispense cases summarily. It can significantly reduce the time and effort needed to analyze a case, and courts can apply consistently evolving legal principles, even when the facts are idiosyncratic.⁴¹¹ It can also avoid the risk of judges engaging in side-by-side mark comparison and ensure they apply the real-world purchasing context. The results from AI recommendations challenge judges' prior assumptions, providing a check against coherence-based reasoning. Simon's research shows that confronting people with merits of the opposite side reduced the effect of coherence shifts by about 50%.⁴¹² In particular, his study moderated jury instruction by expressly requesting jury members to "take some time to seriously consider the possibility that the opposite side has a better case."⁴¹³ Other legal studies similarly showed that asking lawyers to consider the weaknesses in their side or reasons that the judge might rule against them mitigated bias.⁴¹⁴

The beauty of AI-enabled likelihood of confusion analysis is that it can reach outcomes we cannot define in advance of the AI being run as "good" or "better" than the untrained neural network interrogates itself via the process of trial and error. In addition, convolutional neural networks can abstract local features from examples, for instance, by recognizing specific facts in opinions. They would also account for interactions among indicators that escape even expert

408. See Daryl Lim, *Confusion, Simplified*, BERKELEY TECH. L.J. (Forthcoming, 2022) [hereinafter Lim, *Confusion, Simplified*].

409. *Id.*

410. *Id.*

411. Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537, 2541–42 (2009).

412. Simon, *Third View of the Black Box*, *supra* note 225, at 544 (noting that "[m]ore studies are required to gain a better sense of the effects of the debiasing intervention").

413. *Id.* at 571.

414. See Linda Babcock et al., *Creating Convergence: Debiasing Biased Litigants*, 22 L. & SOC. INQUIRY 913, 920–21 (1997).

witnesses and contextualize and associate information with known factors to provide predictions based on untrained parameters.⁴¹⁵ Finally, unsupervised data mining algorithms can zero in on data clusters and probe those clusters to find other abstractions.⁴¹⁶

Moreover, programming the AI to maximize reward in a predetermined environment allows it to directly optimize policy performance rather than learning from old data⁴¹⁷ by updating the agent's policy using good estimates of a particular policy's advantage relative to another policy.⁴¹⁸ Conceivably, variations of the algorithm will predict litigation risk and the business implications of marketing and sales decisions.⁴¹⁹ As Dev Gangjee put it, "it is extremely tempting to be guided by clearly defined percentages of similarity."⁴²⁰ A. S. Li, A.J.C. Trappey, and C.V. Trappey sketched out how that model might work. The data set combines trademark litigation ontology and text mining to extract features from cases to build a machine-readable database like case content analysis.⁴²¹

2. Execution

Like many AI datasets, case content analysis treats the content of opinions as generic data.⁴²² Coding and counting cases imply that

415. Similarly, AI-based support vector machines (SVMs) can find relationships between sets of trademark infringement cases while handling outlier or mislabeled cases, allowing SVM to crunch abrogated case law. *See e.g.*, AURÉLIEN GÉRON, HANDS-ON MACHINE LEARNING WITH SCIKIT-LEARN, KERAS, AND TENSORFLOW 155–67 (Nicole Tache ed., 2d ed. 2019) (explaining how SVMs work and how they can be helpful).

416. PEDRO DOMINGOS, THE MASTER ALGORITHM 210 (2015).

417. *Proximal Policy Optimization, OpenAI Spinning Up*, OPENAI (2018) <https://spinningup.openai.com/en/latest/algorithms/ppo.html> [<https://perma.cc/GB72-ZWGX>].

418. *See* Brian S. Haney, *AI Patents: A Data Driven Approach* 19 CHI-KENT J. INTELL. PROP. 407, 439 (2020) (explaining the advantage function).

419. Katyal & Kesari, *supra* note 404, at 533 ("Indeed, predictive analytics can prove to be transformative in helping businesses both create and sustain a strong presence in the marketplace, predicting the outcome of filing suit, sending a cease-and-desist, articulating various claims, or deciding whether and for how much to settle. And this is just the tip of the iceberg. Imagine every aspect of a trademark claim—its probable outcome automated, calculated, predicted and ready for real-time decision-making.").

420. *See* Gangjee, *supra* note 404, at 13.

421. A. S. Li, A. J. C. Trappey, & C. V. Trappey, *Intelligent Identification of Trademark Case Precedents Using Semantic Ontology*, in *TRANSDISCIPLINARY ENGINEERING FOR COMPLEX SOCIO-TECHNICAL SYSTEMS—REAL-LIFE APPLICATIONS* (Jerzy Pokojski et al. eds. 2020).

422. Hall & Wright, *supra* note 29, at 83.

information in one opinion is potentially relevant to another.⁴²³ In a pre-AI world, an army of legal scholars might attempt to map all likelihood of confusion cases comprehensively. Indeed, they would need to endure many hours of time-consuming and repetitious reading and extracting of the necessary information to code each case, draw inferences, and report on trends, as was done in this Article's writing.

Case content analysis is suitable for automating because the same set of information must be keyed into many cells in the same case. This requires coders to eyeball each cell for accuracy given the tedious, repetitive data entry, resulting in avoidable human errors and copy-paste tasks.⁴²⁴ Nevertheless, as seen in this Article, the result is useful, capable of determining the weight courts have placed on various legal and non-legal factors, identifying which factors judges use to "stampede" others, revealing trends across time, and other relevant parametric factors that may typically escape conventional wisdom.⁴²⁵

The algorithm would pick out keywords and assign appropriate weights to each variable with AI. For example, factors "in favor of"⁴²⁶ or "favors"⁴²⁷ would signify a positive correlation to one side. Similarly, phrases like "marks are strong,"⁴²⁸ "high degree of care,"⁴²⁹ "marks are . . . identical,"⁴³⁰ and "weighs heavily in favor of,"⁴³¹ would be assigned greater, or in the case of "neutral,"⁴³² "weighs neither for

423. See Peter J. Hammer & William M. Sage, *Antitrust, Health Care Quality, and the Courts*, 102 COLUM. L. REV. 545, 561 (2002) (explaining that the coding exercise does not determine the law, but instead treats opinions as data).

424. Serena Lim & Nandini Nayar Sharma, *Document Drafting—Less Is More*, SING. L. GAZETTE (Nov. 21, 2021), <https://lawgazette.com.sg/practice/tech-talk/document-drafting-less-is-more>.

425. See e.g., Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005*, 156 U. PA. L. REV. 549, 591 (2008) (describing a study of cases regarding the fair use reasoning used by different courts).

426. Fletcher's Original State Fair Corny Dogs, LLC v. Fletcher-Warner Holdings LLC, 434 F. Supp. 3d 473, 485 (E.D. Tex. 2020).

427. E.g., Fleet Feet, Inc. v. Nike Inc., 419 F. Supp. 3d 919, 943 (M.D.N.C. 2019) ("This factor favors Fleet Feet."), *appeal dismissed*, 986 F.3d 458 (4th Cir. 2021).

428. New Balance Athletics, Inc. v. USA New Bunren Int'l Co., 424 F. Supp. 3d 334, 347 (D. Del. 2019).

429. *Id.* at 348.

430. *Id.* at 347.

431. Better Angels Soc'y, Inc. v. Inst. For Am. Values, Inc., 419 F. Supp. 3d 765, 777 (S.D.N.Y. 2019); Life After Hate, Inc. v. Free Radicals Project, Inc., 410 F. Supp. 3d 891, 908 (N.D. Ill. 2019).

432. Illinois Tool Works Inc. v. J-B Weld Co., 419 F. Supp. 3d 382, 399 (D. Conn. 2019).

[n]or against,”⁴³³ “slightly in favor of,”⁴³⁴ would be afforded less weight. The algorithm would also recognize and capture variables like rivalry (“direct competitors”⁴³⁵).

Automation saves dataset preparers a substantial amount of time. Studies on automating conveyancing work show a time savings of 90%.⁴³⁶ Less skilled and lower-cost staff can quickly and accurately generate datasets, lowering costs, time, and effort to produce complex datasets, freeing up scholars to focus on higher-value work.⁴³⁷ The user selects a smart template and answers a questionnaire presented by the template to generate an opinion.⁴³⁸ The AI then uses the training data to assemble a custom opinion.⁴³⁹

Likelihood of confusion opinions contain logic-dependent conditional clause variations which incorporate the factors. The algorithm could compare the qualitative and quantitative factors presented in each case to its markers as a first step. Cases presenting the same set of facts would reach the same outcome as precedential cases presenting the same set of markers. AI will need to specify the weight of factors not expressly entailed by rules or precedents. Once algorithms produce their recommendation, judges could accept or reject the AI’s recommendation, like how Amazon consumers choose to make another purchase based on Amazon’s recommendations of their earlier purchases and browsing history.⁴⁴⁰

The algorithm randomly plays out certain results, learns—with input from data scientists in each iteration—adjusts its weights and parameters, and chooses advantageous moves with increasing finesse.⁴⁴¹ The feedback loop causes the algorithm’s nodes to change their weights, so case precedents refined by new case law and market

433. *New Balance Athletics, Inc.*, 424 F. Supp. 3d, at 349.

434. *Delta Forensic Eng’r, Inc. v. Delta V Biomechanics, Inc.*, 402 F. Supp. 3d 902, 910 (C.D. Cal. 2019).

435. *Smartling, Inc. v. Skawa Innovation Ltd.*, 358 F. Supp. 3d 124, 139 (D. Mass. 2019).

436. *Lim & Sharma*, *supra* note 424.

437. *Id.*

438. *Id.*

439. *Id.*

440. See SPANDANA SINGH, NEW AMERICA, WHY AM I SEEING THIS? HOW VIDEO AND E-COMMERCE PLATFORMS USE RECOMMENDATION SYSTEMS TO SHAPE USER EXPERIENCE 22 (Mar. 2020) (discussing how Amazon’s recommendation system drives user purchases and contributes to revenue generation on the platform).

441. See TERRENCE J. SEJNOWSKI, THE DEEP LEARNING REVOLUTION 20 (2018) (describing how the AlphaZero used machine learning algorithm to play chess).

data may eventually yield a different, better outcome over time.⁴⁴² This allows adjudicating to become less a question of “ideology plus facts plus law equal the outcome” and more a question of whether the data supports the parties’ legal outcome or if on appeal, one that the lower court advanced.

3. *Limitations*

As with any AI system, there are limitations, some generic, some specific, that its implementers need to keep in mind. First, this Section identifies the main limitations pertinent to the discussion: “garbage-in, garbage-out,” biases, contextualizing purchasing conditions, and coding challenges. Then, this Section discusses each one in turn.

a. *“Garbage-in, garbage-out”*

First, the saying “garbage-in, garbage-out” applies to the training dataset. The algorithm applies the judge’s expertise through the opinions coded in the training data while minimizing unreliability.⁴⁴³ The case law may be doctrinally flawed but remain good law. Nonetheless, the algorithm can implement the likelihood of confusion factors more consistently than both the human judges who decide the precedential cases in the dataset and the judges applying those precedents. Moreover, judges adjudicating live cases can compare the model’s prediction with the ground truth and adjust the model’s parameters, minimizing the error between these two values over time. As algorithms gain additional knowledge about the probabilities of occurrence, ambiguity disappears, and the choices become clearer.⁴⁴⁴

Scholars and AI service providers agree that AI augments human decision-making and does not displace it.⁴⁴⁵ As LawPanel’s founder put

442. See Stephen McJohn & Ian McJohn, *Fair Use and Machine Learning*, 12 NE. U. L. REV. 99, 135 (2020).

443. Dawes, *supra* note 216, at 575 (“[A] linear model distills underlying policy . . . from otherwise variable behavior (e.g., judgments affected by context effects or extraneous variables).”).

444. See Gary Charness & Dan Levin, *When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect*, 95 AM. ECON. REV. 1300, 1300 (2005) (describing a comparable heuristic form of processing new information).

445. See Gangjee, *supra* note 404, at 11 (“Experience till date therefore suggests that AI algorithms are intended to augment human judgment—to effectively sift through ever increasing volumes of registration data—and not to replace it.”); see also COMPUMARK, ARTIFICIAL INTELLIGENCE, HUMAN EXPERTISE: HOW TECHNOLOGY AND TRADEMARK EXPERTS WORK TOGETHER TO MEET TODAY’S IP CHALLENGES 5 (2018) (“While AI and neural networks will play an expanding role in CompuMark solutions . . . they are intended to complement, not replace, human analysts.”).

it, “AI will speed up legal research, but it will not replace advice formulation . . . [since it] only works on repetitive tasks in a very tightly-defined domain.”⁴⁴⁶ Nevertheless, Katyal and Kesari are optimistic that the gap can be closed as data scientists enrich the dataset with more data points and human-AI teams.⁴⁴⁷ They report how experts are continuing to highlight the need for human oversight and participation, particularly when it comes to complex cognitive tasks in trademark doctrines.⁴⁴⁸

On appeal, the variability of decisions can reveal some idea of the extent of noise. A three-judge circuit appeals court or nine-justice Supreme Court bench would provide an additional check. Salib observes that “there will be an adjustment period as courts develop doctrine about what constitutes credible scientific practice in algorithmic design. Such bumps on the road, however, are the cost of admission if generalist judges are to continue playing any major role governing our increasingly complex world.”⁴⁴⁹

b. Biases

Second, system architects need to address data biases in adopting the technology and in deploying AI. For example, with supervised machine learning, humans classify the data. This introduces bias, such as training an AI on the similarity of signs. One trainer might determine a similarity between two given signs, while another might not.⁴⁵⁰ As a result, AI may replicate and perpetuate data biases.⁴⁵¹

Coding is not value-neutral, and biases may seep into the algorithmic code, filtering into training data and the weights judges assign to the algorithm.⁴⁵² Bias could also come from the algorithms being trained

446. Tim Lince, “No Imminent AI Apocalypse”—Tech Expert Rejects Predictions of Mass Job Losses in Trademark Industry, *WORLD TRADEMARK REV.* (Feb. 1, 2018), <https://www.worldtrademarkreview.com/brand-management/no-imminent-ai-apocalypse-tech-expert-rejects-predictions-mass-job-losses> [<https://perma.cc/RAF4-FHPV>].

447. See Katyal & Kesari, *supra* note 404, at 526.

448. *Id.* at 533.

449. Peter N. Salib, *Artificially Intelligent Class Actions*, 100 *TEX. L. REV.* 519, 552 (2022).

450. See Lim, *Confusion, Simplified*, *supra* note 408.

451. See *id.*

452. See Dan L. Burk, *Algorithmic Fair Use*, 86 *U. CHI. L. REV.* 283, 283 (2019) (describing how design values of algorithms can reflect biases); see, e.g., David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 *U.C. DAVIS L. REV.* 653, 669–701 (2017) (explaining that because humans make algorithms and humans have bias, the biases of humans are reflected in both algorithms themselves and how humans use them).

using biased data, such as prior decisions from judges who are biased themselves, and from the way humans interpret the data produced by AI systems.⁴⁵³ In addition, reinforcement learning techniques may embed bias, raising the risk of what Thomas Nachbar labeled “snowballing unfairness.”⁴⁵⁴ Codes are based on earlier program decisions and the constant integration of new information, prompts a continual search for purpose.⁴⁵⁵

Moerland and Freitas provide an example of bias in action:

[W]hen teaching an AI to establish a pattern of similarity of signs, one could easily ascertain a similarity between two given signs, while someone else would not. Even if case law regarding similarity of signs is used as training data, courts sometimes come to differing outcomes for the same cases. Bias in data will be replicated when used by the AI technology, as it lacks the ability to filter out slightly incorrect interpretations.⁴⁵⁶

The lack of a standardized method to weigh factors systematically exacerbates the risk of bias. The likelihood of confusion factors have no weights assigned, eroding the ability to apply the tests objectively or in a manner that can be replicated.⁴⁵⁷ AI helps integrate data and provides a statistical prediction based on input variables. Humans are superior at selecting and coding information but poor at integrating it.⁴⁵⁸

Daniel Kahneman, Cass Sunstein, and Olivier Sibony recommend assigning probabilities rather than absolute values or binary “yes” or “no” judgments.⁴⁵⁹ Numerical thresholds could help adjudicate infringement cases. For example, computer scientists could build a model that requires judges to rate the three core likelihood of confusion factors on a scale of 0–10. If the marks were completely different, the judge would rate it ‘0’ (the lowest rating possible), but if

453. See Gangjee, *supra* note 404, at 11 (“[W]here the data for a machine learning approach is derived from judicial content analysis—past decisions by human tribunals where factors are coded and correlations derived—the algorithm will behave like the human decision maker it is modelled after, warts and all.”).

454. Thomas B. Nachbar, *Algorithmic Fairness, Algorithmic Discrimination*, 48 FLA. STATE L. REV. 509, 522 (2021).

455. *Id.* at 548.

456. Moerland & Freitas, *supra* note 404, at 282.

457. See *Menard, Inc. v. Comm’r of Internal Revenue*, 560 F.3d 620, 622–23 (7th Cir. 2009) (“Multifactor tests with no weight assigned to any factor are bad enough from the standpoint of providing an objective basis for a judicial decision; multifactor tests when none of the factors is concrete are worse.”).

458. See, e.g., Dawes, *supra* note 216, at 573.

459. See KAHNEMAN, SIBONY & SUNSTEIN, *supra* note 176, at 218.

the marks were simple counterfeits, the judge would rate it ‘10’ (the highest rating). Thus, the algorithm would set a numerical threshold for finding confusion that maps to case law and the balance of probabilities. Over time, the algorithm would provide more granular information about the characteristics driving outcomes in likelihood of confusion cases. In this way, the algorithm would imitate judges, granting a low score to a particular factor and a consequently lower success rate to plaintiffs.

Daniel Kahneman, Cass Sunstein, and Olivier Sibony also recommend relying more heavily on rules like judicial sentencing guidelines.⁴⁶⁰ The trio of factors again provides that framework. Importantly, the results from AI recommendations challenge judges’ prior assumptions, providing a check against coherence-based reasoning.⁴⁶¹ For instance, confronting people with merits of the opposite side reduced the effect of coherence shifts by about 50%.⁴⁶² Legal studies similarly showed that asking lawyers to consider the weaknesses in their side or reasons that the judge might rule against them mitigated bias.⁴⁶³

Finally, to address the issue of “snowballing unfairness,” flooding the system with voluminous data may help. As Moerland and Freitas note, “with large amounts of data, incidental bias may not influence the rule that the AI learns from the data.”⁴⁶⁴ They reassuringly report that “[AI] training is continuous and subject to high standards of reliability. Error measures are used as well as pilot studies on unseen data to determine how the AI tool performs its tasks.”⁴⁶⁵

Done well, trademark algocracy will minimize biases from human decision-making without compounding those biases with its own.⁴⁶⁶ In the years ahead, ethics teams will likely become an essential department in antitrust agencies and economic consultancies such as finance, legal, marketing, and human resource departments. These

460. *Id.* at 258.

461. See generally Lim, *Saving Substantial Similarity*, *supra* note 281, at 634 (discussing the safety valve role of juries in challenging the assumptions of judges).

462. Simon, *Third View of the Black Box*, *supra* note 225, at 544 (noting that “[m]ore studies are required to gain a better sense of the effects of the debiasing intervention”).

463. Babcock et al., *supra* note 414, at 920–21 (describing results from an experiment which showed that when subjects consciously considered weaknesses in their cases, their biases were mitigated).

464. See Moerland & Freitas, *supra* note 404, at 282.

465. *Id.* at 281.

466. See Meghan J. Ryan, *Secret Conviction Programs*, 77 WASH. & LEE L. REV. 269, 286 (2020) (describing how computer programs can effectively help judges avoid injecting their own biases when making judicial decisions).

teams can help decision-makers weigh the benefits and harms of AI procedures and recommendations, flag their implications, develop guidelines, and help clarify ethical conflicts.⁴⁶⁷

c. Contextualizing the purchasing conditions

Third, the algorithm needs to replicate how a human perceives a mark in the marketplace.⁴⁶⁸ As we saw in Part I, a human fact-finder trying to contextualize the marketplace faces a difficult task. AI has the additional burden of delivering on that promise of an objective assessment.⁴⁶⁹ The algorithm will need to account for the relevant consumer type, competitors, circumstances of purchase, and the end-use of the product. The AI tool is unlikely to contextualize and juxtapose them against case law to compare situations and determine a likely outcome. The data scientist will need to acquire this information and structure it in a manner that the algorithm can automatically process.

One way to contextualize purchasing conditions is for AI to maximize a preset reward without the need for continual human supervision. In this case, the reward is whether, on balance, consumers would be confused.⁴⁷⁰ Instead, the algorithm chooses an action in the environment's initial state—representing a moment in time—randomly explores the environment, gathers information about the environment, develops an optimal policy, and optimizes performance by “expressing the relationship between the value of a state and the values of future states.”⁴⁷¹

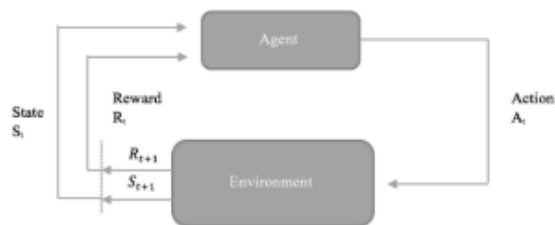
467. Lim Sun & Jeffrey Chan Kok Hui, *Moving AI Ethics Beyond Guidelines*, STRAITS TIMES (Dec. 16, 2020), <https://www.straitstimes.com/opinion/moving-ai-ethics-beyond-guidelines-0> [<https://perma.cc/R4WD-76TD>].

468. Moerland & Freitas, *supra* note 404, at 284 (“This leads us to the finding that the assessment is one of degree and requires reasoning from the perspective of the relevant public. It is questionable as to how far AI technology can reflect this human-centric approach.”).

469. Katyal & Kesari, *supra* note 404, at 532 (“Others have expressed similar concerns, noting that determining trademark distinctiveness, the relevant public, the proper classification of goods and services, among other elements, are so subjective that they pose challenges to the development of AI in trademark law.”).

470. See Haney, *supra* note 418, at 430 (describing how reward can act as a feedback mechanism).

471. *Id.* at 437.

Figure 16: Reinforcement Learning⁴⁷²

As it continues to the next state, the agent receives a reward and a set of choices, the algorithm selects an action, and the environment returns a reward and the next state.⁴⁷³ The reward teaches the algorithm what to do and formalizes the goal's idea.⁴⁷⁴ Through this iteration, it learns to take actions optimizing a reward, which would be, say, mark similarity.⁴⁷⁵ In essence, the total reward mirrors the legal “algorithm” we call trademark law’s likelihood of confusion.

This feature allows the algorithm to navigate dynamic market environments without stopping the environment before computing.⁴⁷⁶ To the extent variables in its dataset need modification, AI training techniques use autoencoders to update word embeddings, machine translation, document clustering, sentiment analysis, and paraphrase detection.⁴⁷⁷ Stacking autoencoders on top of each other allows the first autoencoder to focus on encoding features at one level of abstraction.⁴⁷⁸ The next autoencoder uses the earlier output to recognize fact patterns and encode more abstract features.⁴⁷⁹ Defining features broadly helps avoid overfitting, which happens when the

472. RICHARD S. SUTTON & ANDREW G. BARTO, *REINFORCEMENT LEARNING: AN INTRODUCTION* 48 (2018).

473. EUGENE CHARNIAK, *INTRODUCTION TO DEEP LEARNING* 113 (2018).

474. *Id.*

475. Jennifer Barry, Daniel T. Barry, & Scott Aaronson, *Quantum Partially Observable Markov Decision Processes*, *PHYS. REV.*, No. 032311, 2014, at 1, 2, <https://journals.aps.org/pr/abstract/10.1103/PhysRevA.90.032311>.

476. *Id.*

477. See Venkata Krishna Jonnalagadda, *Sparse, Stacked and Variational Autoencoder*, *MEDIUM* (Dec. 6, 2018), <https://medium.com/@venkatakrishna.jonnalagadda/sparse-stacked-and-variational-autoencoder-efe5bfe73b64> [<https://perma.cc/ZH2D-JTNW>] (describing the various types of and uses for autoencoders).

478. *Id.*

479. Gideon Lewis-Kraus, *The Great A.I. Awakening*, *N.Y. TIMES*, (Dec. 14, 2016), <https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html> [<https://perma.cc/46B3-8DKV>].

learner fits the function to the data.⁴⁸⁰ Overfitting also happens in legal reasoning when one ties a rule to the facts. The solution is to include more training examples and test the function against other test examples.⁴⁸¹

d. Coding challenges

Fourth, the likelihood of confusion factors do not currently lend themselves to easy coding by a machine, given the coherence-based reasoning and non-uniformity in how courts operationalize those factors, as discussed in this Article. Finally, the Eighth Circuit reminds us that “factors do not operate in a mathematically precise formula.”⁴⁸²

Again, the issue is real but not insurmountable. Courts can do their part by employing more rule-like formulations when applying the likelihood of confusion standard, such as the “rules of thumb” advanced in this Article. Courts can also standardize their lexicon, enabling them to present their judicial opinions in a way more amenable to machine learning. Finally, courts could and should also standardize their treatment of absent factors in the likelihood of confusion inquiry—do these factors favor either party and if so, in what way?⁴⁸³ This templating exercise helps rationalize and consolidate disparate variations into a reusable asset that captures and preserves the substantial knowledge of experienced judges.

Finally, the algorithm will need to distinguish between cases from courts at different levels of the judicial hierarchy. *Stare decisis* tells us that Supreme Court cases take precedence over court of appeal cases, which in turn take precedence over district court cases. However, empirical legal studies routinely ignore the weight *stare decisis* endows in coding datasets.⁴⁸⁴ It matters little if the Supreme Court or a district court looked at likelihood of confusion if the variable of interest is competitive proximity. The algorithm will need to consider judicial hierarchy, the appellate jurisdiction of regional circuit courts, and similar factors as appropriate.

480. GÉRON, *supra* note 415, at 26–28.

481. *Id.* at 29.

482. *Frosty Treats, Inc. v. Sony Comput. Ent. Am., Inc.*, 426 F.3d 1001, 1008 (8th Cir. 2005).

483. *Easy Spirit, LLC v. Skechers U.S.A., Inc.*, 515 F. Supp. 3d 47, 74 (S.D.N.Y. 2021) (explaining that absence of actual confusion favors junior user).

484. See e.g., Lim, *Saving Substantial Similarity*, *supra* note 281.

CONCLUSION

Congress built a degree of indeterminacy into the likelihood of confusion standard as a feature and not a bug. Over the years, however, the jurisprudential roots of trademark law became unruly and tangled. Unwanted variability and bias in judgments cause serious problems by including complex and irrelevant factors, including financial loss and rampant unfairness. Meanwhile, simple rules and algorithms have developed with technological strides presenting big advantages over human judges. Three core factors, combined with two safe harbors and AI, would enable courts to reach consistent and accurate results. A simplified framework promotes fair play, safeguards expressive uses, and enhances access to justice.

This Article presented a contemporary empirical analysis of the likelihood of confusion factors and how they interact. Conventional wisdom teaches us that courts should comprehensively traverse each factor and that likelihood of confusion cases generally require jury determination. The data reveals that neither is true. Instead, courts provide early off-ramps to litigants by “economizing” using a handful of factors or by “folding” factors within each other. The findings also indicate which forums are pro-defendant and which are pro-plaintiff, the impact of rivalry and fair use on outcomes, and the Ninth Circuit’s dominance. This Article also showed how AI systems could use empirical studies as training data to help stakeholders make likelihood of confusion analyses. A familiar yet concise, precise, and efficient framework helps preempt, counsel, and adjudicate disputes. In this way, the likelihood of confusion standard can attain the amphibious benefits of becoming more rule-like while retaining its suppleness.