The Missing Element in Trademark Infringement

Sepehr Shahshahani[*] *& Maggie Wittlin*[**]

ABSTRACT: As trademark law is currently litigated and understood, a plaintiff may succeed on an infringement claim by showing that (1) it owns a valid trademark, and (2) the defendant used a mark in commerce that is likely to confuse consumers into thinking that the plaintiff's and defendant's products come from the same source. We argue that this conventional understanding of the cause of action is missing an element: The plaintiff should also be required to show that the confusion arises from protectable features of the plaintiff's trademark. Without this "causal tracing" requirement, plaintiffs can effectively claim exclusive rights to features that the law allows anyone to use, such as functional features or generic terms. Eliding the link between protectable features and confusion can hinder competition and undermine the purposes of trademark law.

Beyond theoretically demonstrating the need to fill this gap in the law, we show how to do it in practice. We focus on consumer surveys—an increasingly critical component of trademark litigation—and explain how to design surveys that address whether the missing element is satisfied. In doing so, we unearth and challenge an implicit assumption in current survey methodology: that a feature of a trademark must be a but-for cause of a consumer's confusion in order for that consumer to count toward the rate of confusion attributable to

[*] Associate Professor, Fordham Law School. sshahshahani@fordham.edu.

[**] Associate Professor, Fordham Law School. maggie.wittlin@fordham.edu.

For helpful comments, we thank Atinuke Adediran, Scott Baker, Stefan Bechtold, Barton Beebe, Doni Bloomfield, TJ Chiang, Kevin Collins, Nestor Davidson, Shahar Dillbary, Jens Frankenreiter, Janet Freilich, Jeannie Fromer, Dan Hemel, James Hicks, Bruce Kobayashi, Christa Laser, Mark Lemley, Gabe Levine, Jake Linford, Bill McGeveran, Mark McKenna, Mike Meurer, Chris Newman, Jacob Noti-Victor, Emma Perot, Kenneth Simons, Neel Sukhatme, David O. Taylor, Josh Teitelbaum, Rebecca Tushnet, Saurabh Vishnubhakat, Joel Waldenford, Felix Wu, Ben Zipursky, and participants in the Harvard/NYU/Penn Trademark and Unfair Competition Scholarship Roundtable, the University of Virginia Law & Economics Workshop, the BU Law & Economics Workshop, the Cardozo Intellectual Property Colloquium, the SMU Dedman School of Law Faculty Forum, the Munich Summer Institute, the Junior Intellectual Property Scholars Association Summer Workshop, the Intellectual Property Scholars Conference, the George Mason Law & Economics Workshop, the Georgetown Law & Economics Workshop, the NYU Tri-State Region IP Workshop, and the WashU Law Faculty Workshop. We also thank Hooman Yazdanian for excellent research assistance.

that feature. We draw on broader debates about causation to argue that sufficient causes should count as causes for purposes of trademark confusion. We then lay out how to design surveys to take account of multiple sufficient causes. Finally, we analyze several sets of hypothetical results for a survey we designed based on a real case, comparing them to results under existing survey methods and showing that our approach is more helpful to courts. In short, we identify a fundamental problem in trademark law, and we offer a concrete, practical solution.

INTRODUCTION

Look at these two energy bars:

Figure 1: Energy Bar Packages[1]

Do you think they were made by the same company? Or that one was licensed by the other? If so, is it because of the overall arrangement of elements on the packaging: the logos and information bars on the left, the strip with the flavor name that goes across to the right side, and the color-block ends? Or is it because both packages have a transparent section that shows the bar? Under trademark law, the maker of the bar on top could not claim the transparency as a trademark, because it's functional—it lets customers see what is in the package.[2] Our intellectual property system leaves the protection of such functional features to patent law, with its limited term of protection and rigorous examination of novelty and inventiveness, not to trademark law, which grants exclusive rights in perpetuity and without such prior examination.[3] And yet, as the law currently functions, Kind, who makes the bar on top, might well prevail on a trademark infringement claim against Clif, who makes the bar on the bottom, if enough consumers believe they were made by the same company just because both have partially transparent wrappers.[4]

1. All images are presented in color in the PDF version of this Article, available on the *Iowa Law Review* website.

2. *See, e.g.*, Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 164 (1995) ("The functionality doctrine prevents trademark law, which seeks to promote competition by protecting a firm's reputation, from instead inhibiting legitimate competition by allowing a producer to control a useful product feature.").

3. *Id.* at 164–65 ("It is the province of patent law, not trademark law, to encourage invention by granting inventors a monopoly over new product designs or functions for a limited time, . . . after which competitors are free to use the innovation. If a product's functional features could be used as trademarks, however, a monopoly over such features could be obtained without regard to whether they qualify as patents and could be extended forever (because trademarks may be renewed in perpetuity).").

4. This example is based on an actual case, *Kind LLC v. Clif Bar & Co.*, No. 14 Civ. 770, 2014 WL 2619817 (S.D.N.Y. June 12, 2014), which we discuss in greater detail below. *See infra* notes 125–35 and accompanying text.

This is a problem—and, in this Article, we show how to fix it. First, we show that many courts' analysis of trademark infringement is deficient. Currently, to succeed on a trademark infringement claim, a plaintiff must show only that (1) it owns a valid trademark, and (2) the defendant used a mark in commerce that is likely to confuse consumers into thinking that the plaintiff's and defendant's products come from the same or affiliated sources.[5] But this two-part structure is missing a critical link between the two elements: It does not require that consumer confusion be caused by those features of the plaintiff's trademark that make that mark valid in the first place. This means that trademark owners can extend their rights beyond what the validity requirements would allow—they can win a lawsuit against a defendant because that defendant's mark had a confusingly similar feature that the law says *anyone* should be able to use (for example, a functional feature like Kind Bar's transparent wrapping).

Allowing plaintiffs to win based on confusion caused by unprotectable features undermines the purposes of trademark law. Trademark law benefits consumers by making it easier for them to find the products they want, and by giving brands a credible way to signal the source of their products, thereby incentivizing them to develop high-quality products they can put their mark on.[6] The Kind Bar packaging as a whole—the color bars, the color-block ends, the transparent material, the word mark in the middle, all taken together—helps consumers identify the bar as one made by Kind. But allowing a mark-holder to succeed against a defendant based on, say, a functional feature would deprive other companies of a useful feature, giving consumers less competition and a lower-quality selection of products.[7] For example, energy bar consumers might associate transparent packaging with Kind because the company was the first to use that feature. But transparent packaging has a valuable function: Consumers may want to see the food they are buying, and companies may want to showcase their food products. If Kind could use trademark law to block other energy bar companies from using transparent packaging, that would harm consumers and frustrate the purposes of the law. The prospect of such unjustified liability is what motivates us to propose the linkage element.

Courts do sometimes appear to recognize that confusion must arise from the protectable features of a plaintiff's mark to be actionable.[8] But they frequently

---

5. *See, e.g.*, Rex Real Est. I, L.P. v. Rex Real Est. Exch., Inc., 80 F.4th 607, 616 (5th Cir. 2023); Wreal, LLC v. Amazon.com, Inc., 38 F.4th 114, 126 (11th Cir. 2022); AWGI, LLC v. Atlas Trucking Co., LLC, 998 F.3d 258, 264 (6th Cir. 2021); 3 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 23:1 (5th ed. 2024).

6. *See, e.g.*, *Qualitex*, 514 U.S. at 163–64; Groeneveld Transp. Efficiency, Inc. v. Lubecore Int'l, Inc., 730 F.3d 494, 511–14 (6th Cir. 2013).

7. *See, e.g.*, TrafFix Devices, Inc. v. Mktg. Displays, Inc., 532 U.S. 23, 32–33 (2001); *Qualitex*, 514 U.S. at 164.

8. *See, e.g.*, Dixi-Cola Lab'ys, Inc. v. Coca-Cola Co., 117 F.2d 352, 361 (4th Cir. 1941) (holding that Coca-Cola cannot stop a competitor from using the word "cola" in its mark); Judson Dunaway Corp. v. Hygienic Prods. Co., 178 F.2d 461, 465–66 (1st Cir. 1949) (refusing to find infringement in a case involving two trademarks for toilet cleaners that both depicted a woman pouring the

miss this point, and the requirement is not reflected explicitly in trademark doctrine.[9] This omission can lead to unwarranted liability in a wide range of cases involving trade dress, symbol marks (logos), and word marks.[10] We argue that it is time to bring the "causal linkage" principle out of the shadows and make it an explicit element of a trademark infringement claim.

Our contribution is not only doctrinal and theoretical; we also demonstrate how, in practice, lawyers and courts can determine whether consumers are confused for the right or wrong reasons. We focus on consumer surveys—an increasingly critical part of trademark litigation,[11] especially in high-stakes and close cases.[12] In a typical survey, an expert locates potential consumers, shows them the defendant's product, and asks questions designed to determine whether they believe the defendant's product is made by or affiliated with the plaintiff.[13] If many survey respondents incorrectly believe the products are affiliated, the survey finds a high rate of confusion, which helps to prove likelihood of confusion.[14] But if the marks in question share both protectable and unprotectable features, a simple finding that many consumers are confused does not show *why* they are confused: It does not show that they are confused by the protectable features.

There are practical ways to improve survey estimates of actionable consumer confusion, accounting for the missing element. Some elite survey experts have already designed instruments to ascertain whether consumer confusion arises from protectable or unprotectable features of the trademark.[15] However, we

---

contents of a can into a toilet bowl, and holding that the plaintiff did not have an exclusive right to represent a woman using a toilet cleaner on toilet-cleaning products). For more cases and discussion, see *infra* Section I.B.

9. *See, e.g.*, Taco Cabana Int'l, Inc. v. Two Pesos, Inc., 932 F.2d 1113, 1117 (5th Cir. 1991) (finding a likelihood of confusion between the trade dresses (décors) of two Mexican restaurants even though much of what they had in common consisted of generic, unprotectable elements such as a "festive and vivid color scheme," patios, and bright awnings and umbrellas), *aff'd*, 505 U.S. 763 (1992); Decatur Fed. Savs. & Loan Ass'n v. Peach State Fed. Savs. & Loan Ass'n, No. C-78-904-A, 1978 WL 21348, at *5–6, *8 (N.D. Ga. Dec. 29, 1978) (enjoining the defendant from using "Savings Spot" as a mark for a savings and loan association on the grounds that it was likely to cause confusion with plaintiff's "Savings Shop" mark for the same services because "[b]oth marks contain 'savings' which is the operative word in inducing identification with the services of the parties"). For further discussion of these and many other cases, see *infra* Section I.B.

10. *See infra* notes 56–81 and accompanying text.

11. *See, e.g.*, Shari Seidman Diamond & David J. Franklyn, *Trademark Surveys: An Undulating Path*, 92 TEX. L. REV. 2029, 2030 (2014) (finding, based on a survey of trademark attorneys, that "trademark surveys often play multiple important roles in the life of a trademark case"); Eric D. DeRosia, *Fixing* Ever-Ready*: Repairing and Standardizing the Traditional Survey Measure of Consumer Confusion*, 53 GA. L. REV. 613, 617 n.7 (2019) (reviewing a variety of estimates and concluding that "in trademark litigation, surveys are commonplace but not universal").

12. Diamond & Franklyn, *supra* note 11, at 2054–55, 2067.

13. *See* 5 MCCARTHY, *supra* note 5, § 32:158 (providing a primer on the use of surveys in trademark law).

14. *Id.*

15. *See infra* notes 120–35 and accompanying text.

argue that even these survey methods are not up to the task of determining whether consumers are confused for the right or wrong reasons.

The critical error in existing best practice is that it ignores a fundamental issue of causation. Specifically, it implicitly assumes—without even considering the question—that the protectable features of the marks must be a *but-for* cause of a survey respondent's confusion for that respondent to be counted toward the rate of confusion. In other words, if a respondent would still be confused if the defendant's mark did *not* include features similar to the protectable features of the plaintiff's mark, that respondent should not count. But we argue that the principles of causation accepted in other areas of law, as well as good trademark policy, demand that we also count respondents for whom the protectable features are a *sufficient*, but not necessary, cause of confusion.[16] In other words, if a consumer would be confused by the marks' similar *protectable* features, standing alone, they should not be excluded from the actionable rate of confusion just because they would *also* be confused by the marks' similar *un*protectable features, standing alone.

To fix this problem, we introduce a novel set of control conditions—modified versions of our survey that allow us to tease out the sources of consumer confusion.[17] Specifically, these controls allow us to identify (1) how many survey respondents are confused only by protectable features that the marks have in common, like the energy bars' overall package configuration; (2) how many are confused only by unprotectable common features, like the transparent wrapping; (3) how many would be confused by the protectable features alone, but would *also* be confused by the unprotectable features alone; (4) how many are confused only if the marks have *both* protectable and unprotectable features in common; and (5) how many are not confused at all.[18]

The best existing measures of confusion include group (1), consumers who are confused by the protectable features alone, and group (4), consumers who are confused by both the protectable and unprotectable features together but not by either of them alone. But they cannot measure and do not include group (3), consumers who would be confused by either the protectable features alone or the unprotectable features alone. However, as discussed, the common protectable features are a *sufficient* cause of confusion for consumers in group (3), so they should count. Our proposed methodology allows us to distinguish these groups more clearly than the best existing practices can, and especially to capture the third, currently ignored group.

After laying out our proposed survey methodology, we present a sample survey and a set of hypothetical results, and we explain how a court should interpret them.[19] By designing actual controls for an actual case, we show that our proposal is workable. It would limit spurious trademark liability but would not be an insurmountable barrier to proving infringement. And our analysis

16. *See infra* Section II.C.
17. *See infra* Section II.B.
18. *See infra* Section II.B.
19. *See infra* Part III.

of sample results demonstrates how our methodology can provide more accurate and useful results than even the best existing methods.

This Article, therefore, makes two contributions to trademark literature and practice. First, by surfacing an overlooked "causal tracing" element, we make a doctrinal and theoretical contribution to the rich literature on the trademark infringement standard.[20] While other scholars have suggested a need to link validity and infringement in intellectual property generally, they have not pinpointed this problem in trademark law.[21] We show that the problem is worse in trademark than in other intellectual property regimes and that trademark can learn from other regimes by importing safeguards against a finding of infringement not linked to protectable features.[22] Second, we operationalize a solution to the problem by showing how to prove the missing element, in the process significantly advancing trademark survey methodology and contributing to a more sophisticated understanding of causation in trademark law.[23]

This Article proceeds in three parts. Part I discusses the existing law and literature on trademark infringement and argues that it is missing a causal tracing element. Part II shows how to prove the missing element by using surveys to tease out whether consumer confusion arises from common protectable features of the parties' products. It also argues that if consumers would be confused either by the common protectable features alone or by the common unprotectable features alone, they should count toward the rate of confusion. Part III demonstrates how our method works by presenting a sample survey with

---

20. *See, e.g.*, Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581, 1600, 1617–19, 1627–29 (2006) (finding, in a pioneering empirical study of the multifactor test for likelihood of confusion, that courts often rely on one or two factors to reach an outcome and "stampede" the other factors to conform to that outcome, and that the "bad faith" factor is "nearly dispositive"); Rebecca Tushnet, *Running the Gamut from A to B: Federal Trademark and False Advertising Law*, 159 U. PA. L. REV. 1305, 1308–09, 1360–73 (2011) (arguing that the trademark-infringement and unfair-advertising jurisprudences should learn from each other, and proposing a requirement that confusion be actionable only when it is material to consumers' purchasing decision); Mark A. Lemley & Mark McKenna, *Irrelevant Confusion*, 62 STAN. L. REV. 413, 414–16, 445–48 (2010) (advocating a materiality requirement that links confusion to consumers' purchasing decision); Robert G. Bone, *Taking the Confusion out of "Likelihood of Confusion": Toward a More Sensible Approach to Trademark Infringement*, 106 NW. U. L. REV. 1307, 1348–78 (2012) (contesting the materiality requirement and calling for different approaches to likelihood of confusion depending on whether the "moral" or "economic" justification of trademark law predominates in the context); Michael Grynberg, *The Consumer's Duty of Care in Trademark Law*, *in* RESEARCH HANDBOOK ON TRADEMARK LAW REFORM 326, 327–31 (Graeme B. Dinwoodie & Mark D. Janis eds., 2021) (arguing that likelihood of confusion is inappropriately conceptualized as a purely descriptive inquiry and urging a consideration of whether consumer confusion is reasonable).

21. *See* Mark A. Lemley & Mark P. McKenna, *Scope*, 57 WM. & MARY L. REV. 2197, 2225–59 (2016) (attacking the disconnect between validity and infringement doctrines in copyright, patent, and trademark). Our intervention is different from Lemley and McKenna's both in that we see the trademark problem as unique (and uniquely acute) and in that we operationalize a solution using surveys. *See infra* notes 94–103 and accompanying text.

22. *See infra* notes 34–48 and accompanying text.

23. *See infra* Section II.B.

hypothetical results and explaining how an expert or a court should interpret those results. The Conclusion is followed by an Appendix that fleshes out some technical issues.

I. THEORIZING TRADEMARK INFRINGEMENT'S MISSING ELEMENT

To prevail on a claim of trademark infringement, a plaintiff must show that it owns a valid trademark and that the defendant has used a mark in commerce that is likely to confuse consumers into thinking that the claimant's and defendant's products come from the same source or affiliated sources.[24] For most practical purposes, this statutory command reduces to showing two things: validity and likelihood of confusion.[25] The twofold structure seems quite sensible. It is the same conceptual structure as in many civil claims—namely, you must show that you have a legally protected interest (a right) and that somebody has stepped on that right. To take a simple example for illustration, a claim of conversion in tort law similarly requires the claimant to show that it owns a piece of property and that somebody else took it (with the relevant intent).[26] It is the thesis of this Article that, notwithstanding its seemingly sensible structure, the action for trademark infringement is missing an element.

A. ARGUMENTS FOR THE MISSING ELEMENT

The missing element, we contend, is a requirement to connect the likelihood of confusion to the protectable features of the trademark. It is not enough that the claimant owns a protectable mark and that the defendant's mark is likely confusing consumers; rather, to be actionable, the confusion must arise from the similarity of the two marks *with respect to the protectable features*. The justification for this "linkage" or "causal tracing" element is that, without it, there may be a finding of infringement without any violation. The claimant could win by simply showing that it has a right and that the defendant has stepped on *something*—perhaps on something adjacent to that right, but not necessarily on the right itself.

To see the point more clearly, it is useful to go back to conversion. Imagine a claim for conversion arising from the taking and logging of timber. If the structure of conversion mirrored the present structure of trademark infringement, it would permit a claimant to prevail by showing that it owns many trees in an area and that the defendant felled and logged some trees from the area—without having to show that the trees the defendant took were the same trees that the claimant owned.

---

24. 15 U.S.C. § 1125(a) (2018).

25. *See supra* note 5 and accompanying text.

26. The oft-cited old-school definition of "conversion" is "a taking with the intent of exercising over the chattel an ownership inconsistent with the real owner's right of possession." Fouldes v. Willoughby (1841) 151 Eng. Rep. 1153, 1157. Modern definitions are consistent. *See, e.g.*, Smith v. Maximum Racing, Inc., 136 S.W.3d 337, 341 (Tex. App. 2004) ("Conversion is the unauthorized and wrongful assumption and exercise of dominion and control over the personal property of another to the exclusion of, or inconsistent with, the owner's rights.").

There is a strong policy rationale for the linkage element: It ensures that trademark rights are not effectively extended beyond what the validity requirements allow. Trademark law, by granting a person the exclusive right to use a mark to identify the source of certain products, both reduces consumers' search costs and encourages investments in product quality.[27] It allows consumers to identify a product's source by simply spotting the trademark, freeing them to devote much less time and energy and money to figuring out the source.[28] And it ensures a producer that its investments in product quality will not be wasted by other producers' effectively pretending to be the source of the high-quality products, which would make consumers distrustful of source information in the marketplace and drive down the quality of products.[29] But this goal of competitive integrity in the marketplace would be subverted if exclusive trademark rights were extended beyond distinctive source identifiers.

For example, if one company had an exclusive right to use the word "pizza" in marketing its pizza (a "generic" term, in trademark parlance), other pizza makers would have a harder time communicating what they are selling, which would increase rather than reduce consumer search costs, without any accompanying public benefit.[30] Likewise, if one company had an exclusive right to put ventilation holes in its pizza boxes (a "functional" feature of the product, in trademark parlance), other manufacturers would have a harder time preventing their pizza from getting soggy in transit, which would hinder competition instead of enhancing information integrity in the marketplace.[31]

---

27. *See, e.g.*, Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 163–64 (1995); Groeneveld Transp. Efficiency, Inc. v. Lubecore Int'l, Inc., 730 F.3d 494, 511–14 (6th Cir. 2013).

28. *Qualitex*, 514 U.S. at 163–64 ("In principle, trademark law, by preventing others from copying a source-identifying mark, 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." (quotation marks omitted)).

29. *Id.* at 164 ("[T]he law helps assure a producer that it (and not an imitating competitor) will reap the financial, reputation-related rewards associated with a desirable product. The law thereby encourages the production of quality products . . . ." (citation and quotation marks omitted)). *See generally* George A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q.J. ECON. 488 (1970) (showing that without the ability to signal quality, such as by a trademark, the quality of products in the market deteriorates).

30. *See, e.g.*, Otokoyama Co. v. Wine of Japan Imp., Inc., 175 F.3d 266, 270 (2d Cir. 1999) (defining "generic" as "a term by which the covered goods or services are designated in the language" and explaining that the exclusion of generic terms from trademark protection "protects the interest of the consuming public in understanding the nature of goods offered for sale, as well as a fair marketplace among competitors by insuring that every provider may refer to his goods as what they are").

31. *See* Bayline Partners v. Weyerhaeuser Co., No. C-93-2828, 1994 WL 286337, at *3 (N.D. Cal. Feb. 17, 1994) (concluding that the overall design of a pizza box with ventilation holes and folding tabs was functional); *see also Qualitex*, 514 U.S. at 164 ("The functionality doctrine prevents trademark law, which seeks to promote competition by protecting a firm's reputation, from instead inhibiting legitimate competition by allowing a producer to control a useful product feature.").

Such generic terms and functional features (among other things) are, therefore, not protected by trademark law.[32]

Crucially, this refusal to protect extends to unprotectable elements that are incorporated into valid marks. For example, if "Juliana's Pizza" is a valid mark, the owner can prevent other pizzerias from using "Juliana's Pizza" or "Juliana's" but cannot prevent them from using the word "pizza" in their brand name. Likewise, a pizzeria may have a valid trademark (trade dress) in the overall look of its pizza boxes, which include ventilation holes, but such a trademark should exclude others from using the overall look of the box, not the functional holes as such. The proposed linkage element, by providing that only the confusion attributable to protectable trademark elements is actionable, ensures that trademark rights are not effectively extended to unprotectable elements such as generic terms or functional features.

As such, the linkage requirement we propose would not require amending the federal trademark statute, the Lanham Act.[33] Rather, because all it does is ensure that trademark protection is not effectively extended to unprotectable features, it is already implicit in the law of trademark validity. But the problem is that in the absence of explicit articulation as an element of trademark infringement, courts often miss the linkage requirement. We will presently illustrate this weakness of trademark jurisprudence, but first, it is useful to compare how trademark's intellectual property cousins deal with the same problem.

Like trademark, copyright and patent infringement claims have a validity-infringement structure.[34] Unlike trademark, though, these regimes explicitly incorporate the crucial element linking infringement and validity. Consider, in turn, utility patents (often simply called patents), design patents, and copyright. Utility patents most clearly require the linking of a violation to the underlying exclusive right. The "claims" in a patent determine the precise boundaries of the exclusive right,[35] and a patent claim is infringed only by a product that covers every single element of a claim (the "all-elements rule").[36] Patent law's requirement of defining the scope of exclusive rights in claims,

32. *See, e.g.*, 15 U.S.C. § 1064(3); *Qualitex*, 514 U.S. at 164; *Otokoyama*, 175 F.3d at 270.

33. *See* 15 U.S.C. §§ 1051–1127.

34. Patent's principal validity requirements are in 35 U.S.C. §§ 101–103, 112, and infringement is defined in § 271. Copyright's minimal validity requirements are in 17 U.S.C. § 102(a), with other validity constraints in §§ 102(b) and 101 ("design of a useful article") and elsewhere. Copyright infringement is defined principally in § 501.

35. *See, e.g.*, Markman v. Westview Instruments, Inc., 517 U.S. 370, 372 (1996) (describing the "claim" as "the portion of the patent document that defines the scope of the patentee's rights"); Merrill v. Yeomans, 94 U.S. 568, 570 (1876) (noting that the claim is "of primary importance, in the effort to ascertain precisely what it is that is patented").

36. *See, e.g.*, Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 40 (1997) (describing the "essential inquiry" in patent infringement as follows: "Does the accused product or process contain elements identical or equivalent to each claimed element of the patented invention?").

coupled with the all-elements rule, ensures that infringement can be established only by reference to a specific, predefined, valid claim.[37]

The patent comparison may seem unfair insofar as the technical, scientific nature of patentable innovations seems to allow a more precise *a priori* delineation of the scope of exclusive rights than is practicable in a field like trademark.[38] But the linkage element also exists in design patents and copyright—which, like trademark, allow for a sort of *gestalt* protection, and where, again like trademark, the infringement inquiry is not as precise or mechanical as in utility patent's "all-elements rule." In both design patents and copyright, infringement is determined by asking whether an ordinary consumer would find the designs or the works substantially similar.[39] Both regimes, however, incorporate safeguards to ensure that only substantial similarity *as to protectable elements* counts.

In design patents, this is accomplished by comparing the patented design with the allegedly infringing design *in the context of the "prior art,"* meaning previously existing designs in the marketplace.[40] The point of this "three-way . . . comparison"—as opposed to a two-way comparison between the allegedly infringed and infringing designs—is to deemphasize features common to the two designs that were also present in the prior art, bringing into relief only those features of the allegedly infringed design that are distinct from the prior art and hence protectable.[41] There can be infringement only if an ordinary observer familiar with the prior art considers the designs to be substantially similar.[42]

---

37. For example, one court held that a patent claim on

> [a] toy comprising an elongated housing [case] having a chamber therein for a liquid [tank], a pump including a piston having an exposed rod [piston rod] and extending rearwardly of said toy facilitating manual operation for building up an appreciable amount of pressure in said chamber for ejecting a stream of liquid therefrom an appreciable distance substantially forwardly of said toy, and means for controlling the ejection

was not infringed by a toy gun with an *external* water tank because it did not cover the element of "a 'chamber *therein* for a liquid.'" Larami Corp. v. Amron, No. CIV. A. 91-6145, 1993 WL 69581, at *4 (E.D. Pa. Mar. 11, 1993) (alterations in original) (emphasis added) (quoting U.S. Patent No. 4,239,129).

38. In academic parlance, patent law largely follows a "peripheral claiming" system requiring that the scope of protection be delineated *ex ante*, not a "central claiming" system where the person seeking protection discloses the thing to be protected or its central features and leaves the task of determining the precise scope of protection to the time of litigation over infringement. *See* Jeanne C. Fromer, *Claiming Intellectual Property*, 76 U. CHI. L. REV. 719, 721 (2009); Dan L. Burk & Mark A. Lemley, *Fence Posts or Sign Posts? Rethinking Patent Claim Construction*, 157 U. PA. L. REV. 1743, 1744–46 (2009).

39. *See, e.g.*, Gorham Mfg. Co. v. White, 81 U.S. (14 Wall.) 511, 526–27 (1871) (design patents); Lotus Dev. Corp. v. Borland Int'l, Inc., 49 F.3d 807, 813 (1st Cir. 1995) (copyright).

40. Egyptian Goddess, Inc. v. Swisa, Inc., 543 F.3d 665, 672, 676–77 (Fed. Cir. 2008) (en banc).

41. *Id.* at 672–78.

42. *See id.* at 681–83.

In copyright, prevailing on an infringement claim[43] requires showing the ownership of a valid copyright, copying, and improper appropriation.[44] Copying is proven by a showing of defendant's access to claimant's copyrighted work and of substantial similarity between the two works.[45] Improper appropriation means that what the defendant copied was copyrightable and there is substantial similarity in copyrightable elements.[46] In other words, a violation requires not just copying from a copyrighted work but copying what is copyrightable from a copyrighted work.[47]

A classic example is that an author might have a valid copyright in a story recounting the love between a boy and a girl from two hostile families, but that copyright does not extend to love stories featuring couples from hostile families writ large, nor to stock elements of that genre, so someone who copies those ideas or elements from the author would not be liable for copyright infringement.[48] This obligation to show that the defendant copied something copyrightable is exactly the element connecting the right to the infringement that trademark lacks. We do not claim that the linkage or causal tracing element is perfectly operationalized in other intellectual property regimes; but at least it exists in a more explicit, articulate way than it does in trademark.

In trademark, the linkage element is not likely to bite in cases of arbitrary or fanciful word marks such as Kodak films or Apple computers, nor in cases of highly distinctive symbol marks such as Adidas's iconic logo (Figure 2). With such marks, what is protectable is such a large and prominent portion of the mark as a whole that, in most imaginable cases, any confusion would be attributable to a competing mark's similarity with the protectable parts of the original mark.

---

43. Copyright confers a bundle of exclusive rights, enumerated in 17 U.S.C. § 106 (and § 106A for "moral rights"). We refer here to the first of these rights, which is the right to reproduce, the essential *copy*right. *See* 17 U.S.C. § 106(1).

44. Different courts have different ways of stating these elements, but they come down to essentially the same thing. *See, e.g.*, Arnstein v. Porter, 154 F.2d 464, 468 (2d Cir. 1946); Skidmore v. Led Zeppelin, 952 F.3d 1051, 1064 (9th Cir. 2020) (en banc).

45. *Arnstein*, 154 F.2d at 468; *Led Zeppelin*, 952 F.3d at 1064.

46. *See, e.g.*, Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 361 (1991) ("Not all copying, however, is copyright infringement. To establish infringement, two elements must be proven: (1) ownership of a valid copyright, and (2) copying of constituent elements of the work that are original."); *Led Zeppelin*, 952 F.3d at 1064 ("[O]nly substantial similarity in protectable expression may constitute actionable copying that results in infringement liability . . . .").

47. *Feist*, 499 U.S. at 361–63; *Led Zeppelin*, 952 F.3d at 1064.

48. *See* Nichols v. Universal Pictures Corp., 45 F.2d 119, 121–22 (2d Cir. 1930).

Figure 2: Adidas Logo

But the missing element does real damage in cases where the plaintiff's and defendant's marks are similar with respect to both protectable and unprotectable features, like the energy bar example with which we began this Article.[49] This problem frequently arises with trade dress protecting the "overall look and feel" of a product.[50] In this context, the truism that "the overall impression" is the criterion and individual features are not to be dissected[51] tends to confuse courts into effectively extending trademark protection to unprotectable features.[52]

We fully recognize, of course, that the "overall look and feel" or "overall impression" of a product or package is trademarkable.[53] And we do not question longstanding law that the composite whole may be protectable even if constituent parts standing alone are not.[54] But our point is that to prevail on a claim that the overall look has been infringed, the claimant must show that consumers are likely to be confused *by the overall look*. Simply to show that the overall look is protected and that consumers are likely to be confused leaves open the possibility that they are confused by an unprotectable feature—like the transparent wrapping of an energy bar package or holes in a pizza

49. *See supra* Figure 1 and accompanying text.

50. In line with standard usage and 15 U.S.C. § 1127, we use the word "trademark" to denote not just word marks and logos but the broader category of source identifiers that includes, among other things, trade dress. "Trade dress" is not defined in the Lanham Act. The following useful working definition, cobbled together from various sources, was given by the Supreme Court: "The 'trade dress' of a product is essentially its total image and overall appearance. It involves the total image of a product and may include features such as size, shape, color or color combinations, texture, graphics, or even particular sales techniques." Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763, 764 n.1 (1992) (citation and quotation marks omitted).

51. *See, e.g.*, Exxon Corp. v. Tex. Motor Exch. of Hous., Inc., 628 F.2d 500, 505 (5th Cir. 1980); Est. of P.D. Beckwith, Inc. v. Comm'r of Pats., 252 U.S. 538, 545–46 (1920); 3 MCCARTHY, *supra* note 5, § 23:41 (stating "the 'anti-dissection' rule" and collecting cases).

52. See *infra* Section I.B for many examples.

53. *See Two Pesos*, 505 U.S. at 764 n.1, 769–70 (defining "trade dress" as a product's "total image and overall appearance" and noting that trade dress can be trademarkable).

54. *See, e.g.*, Fuddruckers, Inc. v. Doc's B.R. Others, Inc., 826 F.2d 837, 844 (9th Cir. 1987); Jeffrey Milstein, Inc. v. Greger, Lawlor, Roth, Inc., 58 F.3d 27, 32 (2d Cir. 1995). Indeed, even the most distinctive trademarks, such as the Adidas logo, are often arrangements of individually unprotectable shapes (lines and curves).

box—in which case a finding of liability is, in effect, a decision to protect the unprotectable.[55]

B. How Courts Fail to Enforce the Missing Element

Courts frequently fail to recognize this missing element. For example, in "overall look and feel" trade dress cases, courts routinely allow a finding of liability to be premised on a showing that (1) the overall look is protected, and (2) consumers are likely to be confused, without asking whether consumer confusion arises from the overall look.

A case in point is *Two Pesos v. Taco Cabana*.[56] The case is famous for the Supreme Court's holding that trade dress can be "inherently distinctive"[57] (a holding later narrowed to the product packaging variety of trade dress[58]), but for our purposes, the lower court opinion is notable for missing the causal tracing element. Taco Cabana operated a Mexican restaurant with a trade dress that it described as "a festive eating atmosphere having interior dining and patio areas decorated with artifacts, bright colors, paintings and murals," with an exterior in "a festive and vivid color scheme using top border paint and neon stripes," as well as "[b]right awnings and umbrellas."[59]

---

55. Although our intervention is the idea of a missing *element*, we are not wedded to a particular burden-placing procedure. That is, we are agnostic as to whether it should be the plaintiff's burden, in showing likelihood of confusion, to trace the confusion to protectable elements or whether it should be up to the defendant, initially, to point out an unprotectable element that might be the cause of confusion, at which point the burden would shift to the plaintiff to show that the confusion is in fact due to a protectable element. *Cf.* Demoret v. Zegarelli, 451 F.3d 140, 151 (2d Cir. 2006) (setting out the *McDonnell Douglas* burden-shifting framework, where after the plaintiff makes a prima facie case of discrimination, the burden shifts to the defendant to provide a "legitimate, non-discriminatory reason for the action," at which point the burden shifts back to the plaintiff to prove discrimination); Club Retro, L.L.C. v. Hilton, 568 F.3d 181, 194 (5th Cir. 2009) ("When a defendant invokes qualified immunity, the burden shifts to the plaintiff to demonstrate the inapplicability of the defense."); Spear v. Atrium Med. Corp., 621 F. Supp. 3d 553, 560 (E.D. Pa. 2022) ("After a defendant has raised such a challenge to personal jurisdiction, the burden shifts to the plaintiff to establish the court's jurisdiction over that defendant."). Both our theoretical intervention and our survey-based method of tracing causation go through regardless of the burden-placing framework. And as a practical matter, whatever that framework might be, the defendant would naturally raise the issue in all cases where confusion may plausibly be due to unprotectable elements.

56. Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763 (1992).

57. *Id.* at 776.

58. Wal-Mart Stores, Inc. v. Samara Bros., 529 U.S. 205, 215 (2000).

59. *Two Pesos*, 505 U.S. at 765 (quoting Taco Cabana Int'l, Inc. v. Two Pesos, Inc., 932 F.2d 1113, 1117 (5th Cir. 1991)).

It is clear from this description (and available pictures[60]) that the trade dress includes many generic[61] or stock elements of a Mexican restaurant of the kind—a "festive eating atmosphere," "bright colors," patios, "paintings and murals," "[b]right awnings and umbrellas"—that are no more protectable as trademarks than are tacos.[62] And yet, the jury instructions allowed for a finding of liability—and the jury did find liability—based on a general inquiry into the likelihood of confusion, without tracing the confusion to what is protectable in the trade dress.[63] Interestingly, the Fifth Circuit accepted Two Pesos's argument that Taco Cabana did not have a right to "preclude competitors from using a Mexican theme for a Mexican restaurant";[64] however, it apparently understood that argument as going only to validity, not infringement, so it thought that any concerns about anticompetitive effects were adequately addressed by the trial court's instructions that functional features are not protectable.[65] In thus failing to connect infringement to validity, the court effectively permitted liability to be based on confusion over unprotectable elements.

Another case involving restaurant décor was brought by Fuddruckers against a Phoenix-area restaurant.[66] Fuddruckers's trade dress encompassed a huge variety of visual and nonvisual features relating to its "decor, menu, layout and style of service,"[67] some of which were undoubtedly functional or generic and therefore unprotectable.[68] The Ninth Circuit, though acknowledging

60. See photographs of Taco Cabana and Two Pesos exteriors, in *Two Pesos v. Cabana*, CORNELL U. L. SCH., SOC. SCI. & L., https://courses2.cit.cornell.edu/sociallaw/student_projects/Tradedresspage2.html [https://perma.cc/RKC2-XT6Z]; *Two Pesos v. Taco Cabana*, U. MICH., https://websites.umich.edu/~jdlitman/classes/SI519/TwoPesosvTacoCabana.html [https://perma.cc/7F6C-F7VJ].

61. The term "generic" in the trade dress context has a slightly different, though related, meaning than in the word mark context. It means not the product's name but rather a standard, common, ordinary, or stock feature of the product, or a feature defined at a general level. *See* 1 MCCARTHY, *supra* note 5, § 8:6.50 ("As applied to trade dress, 'generic' is used as a synonym for 'common' or 'ordinary.'"); *see, e.g.*, Jeffrey Milstein, Inc. v. Greger, Lawlor, Roth, Inc., 58 F.3d 27, 33, 34 (2d Cir. 1995) (holding that "a generalized idea" such as plaintiff's idea to "apply the die-cutting technique to photographic greeting cards" is "generic" and unprotectable); Kendall-Jackson Winery, Ltd. v. E. & J. Gallo Winery, 150 F.3d 1042, 1048 (9th Cir. 1998) ("Grape-leaf designs have become generic emblems for wine. Thus, they are not protectable as trademarks."); Yurman Design, Inc. v. PAJ, Inc., 262 F.3d 101, 118 (2d Cir. 2001) ("[T]he fact that a trade dress is composed entirely of commonly used or functional elements might suggest that the dress should be regarded as unprotectable or 'generic' . . . .").

62. *See Taco Cabana*, 932 F.2d at 1117.

63. *See id.*; *Two Pesos*, 505 U.S. at 766.

64. *Taco Cabana*, 932 F.2d at 1118.

65. *Id.* at 1119. The Fifth Circuit also approved of the trial court's instructions that functionality is assessed not as to "individual elements" but as to "the whole collection of elements taken together," *id.*, and of the trial court's similar instructions about distinctiveness. *See id.* at 1119–21.

66. Fuddruckers, Inc. v. Doc's B.R. Others, Inc., 826 F.2d 837, 844 (9th Cir. 1987).

67. *Id.* at 841.

68. These included: food preparation areas visible to customers, food items presented in glass displays, crocks containing cheese and condiments open to patrons, salt and pepper in

the prevalence of unprotectable elements in the trade dress,[69] called for jury instructions inquiring into the likelihood of confusion between the two restaurants without mentioning the need to trace the confusion to protectable features.[70]

Like trade dress cases, cases involving logos or symbol marks often neglect the missing element.[71] So do cases involving compound word marks, where courts often find a likelihood of confusion without regard to whether the confusion might be due to unprotectable (e.g., generic) parts of the marks.[72]

"institutional-sized containers," "neon signs," "many mirrors," the use of a "ceiling music system" to call patrons with ready orders, the offering of bones to customers' dogs, video games outside restrooms, large ceiling lamps, potted floor plants, and ceiling fans. *Id.* at 839–40, 840 n.1.

69. *See id.* at 844 ("Fuddruckers claims trade dress protection for the impression created by a collection of common or functional elements of restaurant decor.").

70. *See id.* at 845.

71. *See, e.g.*, Planters Nut & Chocolate Co. v. Crown Nut Co., 305 F.2d 916, 919–20 (C.C.P.A. 1962) (finding a likelihood of confusion between two "humanized peanut" characters used as trademarks for peanuts and associated products, without any attempt at sifting whether the confusion was due to the concept of using a humanized peanut as a trademark for peanuts (not actionable) or to any similarities between the distinctive features of the two humanized-peanut characters (actionable)); *In re* Triple R Mfg. Corp., 168 U.S.P.Q. (BL) 447, 448 (T.T.A.B. 1970) (finding a likelihood of confusion between two trademarks for automobile oil filters that consisted of a "humanized configuration of an oil drop" without any effort to determine whether the confusion was traceable to protectable elements rather than to the generic idea of using an oil drop as a trademark for oil filters). In some of these cases, the court is mistaken not just about the missing element but also about what features are protectable as a trademark. For example, one court thought that the general idea of humanizing one's product is trademarkable. *Compare Planters Nut*, 305 F.2d at 920 ("[T]he right to illustrate the product is not the right to 'humanize' it to form a symbol to indicate origin. Engaging in the nut business does not entitle one to adopt as a trademark a humanized version of the identical nut which has already been humanized and adopted as a trademark, *and made famous* by the advertising expenditures of another." (emphasis added)), *with id.* at 927 (Worley, C.J., dissenting) ("Humanizing one's wares is as old as commerce itself and I had always thought was, or at least should be, free for all to do.").

72. *See, e.g.*, Crown Radio Corp. v. Soundscriber Corp., 506 F.2d 1392, 1393 (C.C.P.A. 1974) (finding a likelihood of confusion between Soundscriber and Crownscriber as marks for tape recorders and dictation equipment, and noting the identity of the "scriber" portion of the marks, even though the latter is a generic, or at best descriptive, identifier of the kind of product at issue); Coca-Cola Co. v. Clay, 324 F.2d 198, 199–200 (C.C.P.A. 1963) (finding likelihood of confusion between Coca-Cola and Cup-O'-Cola and specifically rejecting the argument that "since 'Cola' is the common descriptive name of a soft drink flavoring, likelihood of confusion or mistake cannot be predicated on the common use of this word in the marks," reasoning that the marks "must be compared as they are used" and "in their entireties"); *In re* Strathmore Prods., Inc., 171 U.S.P.Q. (BL) 766, 768 (T.T.A.B. 1971) (refusing to register "Glisten," appearing in a teardrop shape with a star for the dot in *i* and a star above the word, for furniture polish on the grounds that it is likely to cause confusion with "Gliss'n," previously registered for household cleaning material, even though *to glisten* is a generic or descriptive term for the end result of polishing or cleaning surfaces); Spice Islands, Inc. v. Frank Tea & Spice Co., 505 F.2d 1293, 1295 (C.C.P.A. 1974) (holding that "Spice Tree" should not be registered as a mark for spices because it is likely to cause confusion with "Spice Islands," reasoning that "[o]f paramount interest is not the descriptive [read generic] nature of SPICE, but the overall commercial impression derived by viewing the marks in their entireties in determining whether a likelihood of confusion exists"); Centraz Indus., Inc. v. Spartan

For example, a court enjoined the use of "Savings Spot" as a mark for a savings and loan association on the grounds that it was likely to cause confusion with plaintiff's "Savings Shop" mark for the same services.[73] The case is an egregious example of the missing element in that the court first noted that the constituent words of the plaintiff's mark were unprotectable standing alone[74] but then went on to find confusing similarity based largely on the common word "savings," determining that the second words in the two marks are unimportant.[75] The effect is that the plaintiff got a monopoly on "savings" (or at least on two-word phrases containing "savings") as a mark for a savings and loan association. This comes close to our example of Juliana's Pizza preventing others from using the word "pizza" for their pizza.

Making a similar mistake, the Federal Circuit held that the mark "Fundough" for water-based modeling compounds used as children's toys should not be registered because it is likely to cause confusion with the famous mark "Play-Doh."[76] In analyzing whether consumers would "receive the same commercial impression from both marks," the court simply analyzed the two components of the marks separately: "Fun" is confusingly similar to "play," and "dough" is confusingly similar to "doh," so the marks are confusingly similar.[77] But each of those individual components is likely unprotectable: "Play" and "fun" both denote a common characteristic of children's toys, and trademark law should not afford any toy company a monopoly over all words signifying recreation.[78] And "dough" seems to be a generic term for a water-based modeling compound to play with.[79] Therefore, actionable confusion could

---

Chem. Co., 77 U.S.P.Q.2d (BL) 1698, 1699 (T.T.A.B. 2006) (finding a likelihood of confusion between "iShine" and "Ice Shine" as marks for "floor finishing preparations" and emphasizing the common word "shine," though that word is at best descriptive in this connection).

73. Decatur Fed. Savs. & Loan Ass'n v. Peach State Fed. Savs. & Loan Ass'n, No. C-78-904-A, 1978 WL 21348, at *4–6 (N.D. Ga. Dec. 29, 1978).

74. *See id.* at *2 (finding that "Savings Shop" is protectable but noting that "each separate word is probably descriptive or generic").

75. *See id.* at *6 ("Both marks contain 'savings' which is the operative word in inducing identification with the services of the parties. The other words, 'shop' or 'spot,' are sufficiently unrelated to the savings and loan business in general so as to be less important from the standpoint of memory. That is, there is nothing about them that grabs the mind and intrinsically requires an association with one or the other of the parties. Therefore, the average consumer may recall from the plaintiff's ads that there is a small branch with extended hours of operation called a 'savings ———.'")

76. Kenner Parker Toys Inc. v. Rose Art Indus., Inc., 963 F.2d 350, 355–56 (Fed. Cir. 1992).

77. *Id.* at 354–55.

78. *See* KP Permanent Make-Up, Inc. v. Lasting Impression I, Inc., 543 U.S. 111, 122 (2004) (explaining that the common law and the Lanham Act do not "allow[] anyone to obtain a complete monopoly on use of a descriptive term simply by grabbing it first" and were not "meant to deprive commercial speakers of the ordinary utility of descriptive words"); Am. Cyanamid Corp. v. Connaught Lab'ys, Inc., 800 F.2d 306, 308 (2d Cir. 1986) ("A trademark holder cannot appropriate generic or descriptive terms for its exclusive use.").

79. *Kenner Parker Toys*, 963 F.2d at 355 n.2. The court sidestepped this question because "it [was] unclear that this issue properly arose before the Board." *Id.*

arise only from the *juxtaposition* of the two components. The court failed entirely to address that issue.[80] This is symptomatic of a general problem in the caselaw: With word marks, as with trade dress, courts often imagine that any concern with confusion caused by similarity with an allegedly infringed mark's unprotectable features is obviated by the fact that the mark is valid, thus missing the causal tracing element and failing to connect validity to infringement.[81]

Our point is not that all these cases reached the wrong result but rather that their analysis was lacking. It is possible that in some of the cases, there was a likelihood of confusion traceable to something protectable in the combination of unprotectable elements or of protectable and unprotectable elements; but the problem is that the courts were content to find a likelihood of confusion and made no effort to trace its source.

We do not mean to imply that no court has ever recognized the missing element. Courts do sometimes get it right. For example, there are numerous cases holding that Coca-Cola cannot prevent competitors from using the word "cola" for their carbonated soft drinks.[82] Similarly, the Seventh Circuit held more than a century ago that there was no actionable likelihood of confusion

---

80. Making the same error as *Kenner Parker*'s analysis of "fun" and "play," other courts, including the Federal Circuit, sometimes find likelihood of confusion due to similarity in "meaning" or "impression" when that meaning or impression simply denotes the nature or a common characteristic of the product and is thus unprotectable. *See, e.g.*, *In re* JS ADL, LLC, 777 F. App'x 991, 995 (Fed. Cir. 2019) (finding actionable likelihood of confusion because "ARTISAN*NY and the cited mark ARTESANO NEW YORK CITY, considered in their entireties, each suggest clothing made by skilled tradespersons in New York"); Wrenn v. Boy Scouts of Am., No. C 03-04057, 2008 WL 4792683, at *6 (N.D. Cal. Oct. 28, 2008) (finding that YOUTHSCOUTS infringed BOY SCOUTS because "[t]he contested composite marks both contain words indicating the programs are aimed at children and the second term of each mark is the word 'scout'"); Bose Corp. v. QSC Audio Prods., Inc., 293 F.3d 1367, 1378 (Fed. Cir. 2002) (finding that POWERWAVE is confusingly similar to ACOUSTIC WAVE and WAVE for sound systems and amplifiers because "[t]he presence of the root element WAVE . . . introduces a strong similarity in all three marks" and "carries a strong connotation of sound waves, [similar to] . . . ACOUSTIC WAVE and WAVE"); *In re* Nationwide Indus., Inc., 6 U.S.P.Q.2d (BL) 1882, 1884 (T.T.A.B. 1988) (holding, in the context of marks for rust-penetrating oils, that "where both marks, when applied to the goods in question, are likely to be perceived by purchasers as signifying that the product sold thereunder busts through, or breaks up, rust, . . . the marks create substantially similar commercial impressions, and . . . there is a likelihood of confusion" (footnote omitted)); Burger Chef Sys., Inc. v. Burger Man, Inc., 492 F.2d 1398, 1399 (C.C.P.A. 1974) (finding a likelihood of confusion between BURGER CHEF and BURGER MAN for drive-in burger joints because the man in BURGER MAN's logo was wearing "a chef's hat of sorts," which "conjure[d] up the concept of 'chef'"); Watercare Corp. v. Midwesco-Enter., Inc., 171 U.S.P.Q. (BL) 696, 701 (T.T.A.B. 1971) ("'AQUA-CARE' and 'WATERCARE' [for water-treatment equipment and services] engender the identical commercial impression.").

81. *See, e.g.*, Clamp Mfg. Co. v. Enco Mfg. Co., 870 F.2d 512, 518 (9th Cir. 1989) (holding that NO-TWIST is confusingly similar to KANT-TWIST for clamps because "the two terms convey the same idea or meaning," and that descriptiveness is not a problem because KANT-TWIST is incontestable and therefore cannot be invalidated for being merely descriptive); *see also supra* notes 64–65 and accompanying text (explaining the same mistake in *Two Pesos*).

82. *See, e.g.*, Dixi-Cola Lab'ys, Inc. v. Coca-Cola Co., 117 F.2d 352, 361 (4th Cir. 1941); Coca-Cola Co. v. Snow Crest Beverages, Inc., 162 F.2d 280, 283 (1st Cir. 1947). *But see* Coca-Cola Co. v. Clay, 324 F.2d 198, 199–200 (C.C.P.A. 1963).

between the plaintiff's Sal-Vet and the defendant's SalTone, both being marks for animal-disease remedies whose principal ingredient was salt.[83] The hyphenated compound word "Sal-Vet" might be trademarkable, reasoned the court, "yet where only the common and nonexclusive feature [i.e., 'sal' for 'salt'] is used, there is no infringement."[84]

There are also cases involving logos, pictures, or symbol marks where the courts don't miss the missing element. For example, in a case involving two trademarks for toilet cleaners that depicted a woman pouring the contents of a can into a toilet bowl, the court found that the plaintiff's trademark was valid, but there was no actionable likelihood of confusion because there were many differences between the women depicted in the marks[85] and the plaintiff did not have an exclusive right to represent a woman using a toilet cleaner on toilet-cleaning products.[86] "[T]o quote from the current musical play 'South Pacific', 'There is nothing like the frame of a dame', and we cannot see why the defendant is not as much entitled to take commercial advantage of pulchritude as the plaintiff."[87] Likewise, in finding no actionable likelihood of confusion between two trademarks for bathing suits that depicted girls in bathing suits diving or about to dive into water, a court stressed that the pictures were different and held that the plaintiff's mark, though valid, "does not . . . preclude others manufacturing similar goods from illustrating the same upon girls in a diving, swimming, or any other general position in which a girl clad in a swimming suit would ordinarily be seen."[88] There are also more recent cases showing courts' attentiveness to the missing element.[89]

---

83. S. R. Feil Co. v. John E. Robbins Co., 220 F. 650, 651 (7th Cir. 1915).

84. *Id.* at 652.

85. Judson Dunaway Corp. v. Hygienic Prods. Co., 178 F.2d 461, 466 (1st Cir. 1949) ("The plaintiff's earlier marks showed women dressed as housemaids, and its later ones showed a woman dressed conservatively as a matronly middle-aged housewife. The defendant, on the other hand, never used a housemaid in its advertising, but showed therein a very slim young housewife, perhaps a bride, ultra-fashionably dressed, i.e. in a very short skirt, wearing an elaborate apron, and standing in an exaggeratedly graceful bending posture pouring into a watercloset bowl with an expression of almost ecstatic delight upon her face. This contrasts sharply with the more conservative dress, staid posture, and expression of restrained joy on the faces of the women in all the plaintiff's trade marks, to say nothing of the rather buxom figure of the woman appearing in its later ones.").

86. *See id.* at 465.

87. *Id.* at 466.

88. Jantzen Knitting Mills v. W. Coast Knitting Mills, 46 F.2d 182, 185 (C.C.P.A. 1931). The dissent is notable for making the classic mistake that courts make in missing the missing element—namely, thinking that a mark's validity obviates any concern with trademarking of unprotectable features. *See id.* (Bland, J., dissenting) ("If the diving girl trade-mark of appellant is a valid one and entitled to registration, it would seem that appellee's trade-mark should not be registered for use on goods of the same descriptive properties, since the trade-marks so nearly resemble each other as to be likely to cause confusion or mistake in the mind of the public and will deceive purchasers."); *see also supra* notes 64–65, 81 and accompanying text (demonstrating the same mistake in other cases).

89. *See* Jackpocket, Inc. v. Lottomatrix NY LLC, No. 23-12-cv, 2024 WL 1152520, at *5 & n.2 (2d Cir. Mar. 18, 2024) (holding that the similarity between the marks Jackpocket and Jackpot.com

Such cases, however, are not sufficient in cementing the causal tracing element. For one thing, they consider the missing element only implicitly, not in an explicit or articulate way. They recognize that a senior user cannot prevent others from using an unprotectable feature, but they do not specifically consider whether any likelihood of confusion is traceable to protectable or unprotectable features. What is more, the cases that courts get right are generally easy cases—cases where an unprotectable element is the *only* commonality between the trademarks in suit.[90] The more complex and interesting cases are those where the trademarks share *both* protectable *and* unprotectable elements. In those cases, sifting the source of confusion demands much more careful analysis, and our proposed method provides guidelines for that.[91] More generally, the failure to articulate causal tracing as a requirement, even in cases that get it right, has meant that this essential concept has not penetrated and pervaded the jurisprudence.[92] The result is that the element is either entirely missing or lurking unnamed in the shadows. We hope that our work to bring it into the open will help make it stick.

C. SCHOLARS' INCOMPLETE ENGAGEMENT WITH THE MISSING ELEMENT

Our work is in dialogue with a rich literature critiquing the standard of trademark infringement, particularly the likelihood of confusion element.[93] The piece of scholarship in this literature that comes closest to addressing the

for online lottery services is not actionable because it goes to the unprotectable term "jackpot," and noting that confusion resulting from the common use of "jackpot" "does not weigh in [the plaintiff's] favor"); RiseandShine Corp. v. PepsiCo, Inc., 41 F.4th 112, 124–25 (2d Cir. 2022) (holding that MTN DEW RISE ENERGY as a mark for canned energy drinks does not infringe RISE as a mark for canned cold-brewed coffee because the only similarity between the marks is "Rise," which "is not distinctive" and merely serves "to signify a virtue of the product"); Am. Cyanamid Corp. v. Connaught Lab'ys, Inc., 800 F.2d 306, 308 (2d Cir. 1986) (holding that HibVAX does not infringe HIB-IMUNE as a mark for a vaccine against Haemophilus influenzae type b diseases ("Hib") because "[a] trademark holder cannot appropriate generic or descriptive terms for its exclusive use, and a trademark infringement finding thus cannot be based on the use of a generic or descriptive term such as 'Hib,'" and further because any similarity between the suffixes VAX and IMUNE is merely descriptive of the product's use and therefore not actionable).

90. *See also* 3 MCCARTHY, *supra* note 5, §§ 23:48–23:49 (collecting more cases of this kind).

91. *See infra* Section II.B.

92. Trademark Office practice has not helped remedy this weakness in caselaw. The Trademark Office has statutory authorization to require *disclaimer* of unprotectable components of a mark. 15 U.S.C. § 1056(a). This authority, however, has not translated into any systematic attempt in trademark infringement actions to ensure that confusion is traceable to valid parts of a mark. To begin, the Trademark Office's statutory power to require a disclaimer is, by its terms, discretionary, and the Office has not exercised the discretion in a consistent way; for example, there is no telling when the Office will find a mark "unitary" such that its component parts cannot be separated and disclaimed. *See* 2 MCCARTHY, *supra* note 5, § 19:66. More importantly, Trademark Office disclaimers do not help with the missing causal tracing element because they do not affect the likelihood of confusion inquiry. *Id.* § 19:72 ("A disclaimer is irrelevant in determining the issue of likelihood of confusion.").

93. *See* sources cited *supra* note 20.

problem we identify is Mark Lemley and Mark McKenna's *Scope*.[94] In that article, the authors attack the disconnect between validity and infringement doctrines in all classic intellectual property regimes. A lamentable consequence of this disconnect, they argue, is that "IP owners can and regularly do seek to expand the scope of their IP rights beyond permissible bounds."[95] This leaves courts with two bad options: overreact by invalidating the right altogether or, the road more frequently taken, impose unjustifiable liability by "allow[ing] an IP owner to claim something that the law should not—and indeed purports not to—give her."[96] Lemley and McKenna excoriate not only trademark but also copyright, utility patents, and design patents for embodying a validity-infringement disconnect.[97] And they seek a solution to it in a "single, integrated" evaluation of an IP right's scope.[98]

Lemley and McKenna's diagnosis of the problem, as it pertains to trademarks, shares with ours a concern that liability may be imposed without a specific finding that the defendant stepped on a protectable aspect of a trademark.[99] But our approach to the problem and its solution diverges from theirs in important ways both theoretical and operational. To begin, we disagree with Lemley and McKenna's lumping together of copyright and trademark as both embodying a validity-infringement disconnect. Rather, as discussed above, we think copyright doctrine has the linkage element that trademark is missing, and we think trademark can learn from the conceptual structure of copyright (as well as patent and design patent) in linking the violation to the right.[100]

More significantly, our solution for teasing out the source of confusion with the aid of surveys would be required regardless of whether one accepts Lemley and McKenna's proposal to add "a single integrated scope proceeding" as a new stage of litigation.[101] *Some* determination of the scope of the right, some determination of what is and is not protectable, is required to carry out our intervention—as it would be to carry out any other survey or non-survey method of analyzing whether consumer confusion stems from protectable elements. In many cases, this determination may not require a separate "scope proceeding"; it is black letter law that the protectable elements of a trade dress

94. *See generally* Lemley & McKenna, *supra* note 21.

95. *Id.* at 2267.

96. *Id.* at 2266–67.

97. *Id.* at 2226–66.

98. *Id.* at 2203.

99. *Cf. id.* at 2201 ("[I]n trademark, a party can be deemed infringing because its products look too similar to the plaintiff's and therefore make confusion likely, even if that confusion is attributable to non-source-designating features of the design."); *id.* at 2244 (complaining that "none of the factors explicitly seek to ascertain whether the features of the defendant's mark that are causing the confusion are features that define the plaintiff 's protectable interest").

100. *See supra* notes 43–47 and accompanying text. We agree with Lemley and McKenna, though, that courts sometimes get it wrong in copyright cases.

101. *See* Lemley & McKenna, *supra* note 21, *passim*.

must be clearly identified,[102] and that identification, or an agreement between the parties as to what is protectable, may suffice. In other cases, determining the scope of protection may be more difficult, and a scope proceeding might be useful. For example, what is protectable about the trademark "Seattle's Best Coffee"?[103] Is it the combination of a city, a superlative, and "coffee"? That combination in that order? That combination in that order, but only if the superlative is "best"? Some other property of the combination? That is a difficult question, but it needs to be answered to determine whether any consumer confusion with a similar mark—say, "Best Portland Coffee"—stems from common protectable elements. In that circumstance, our approach and the Lemley–McKenna approach are complementary. Theirs (or some other approach) can be used to determine the scope; ours must be used to determine the source of confusion once the scope is determined.

Finally, another advantage of our solution is that it is concrete. By being precise about the proper design of surveys and offering concrete experimental illustrations, we hope to go beyond positing a theoretically elegant account to offer an operational solution. Our ambition is to tackle a fundamental theoretical problem with a concrete, practical solution.

II. PROVING THE MISSING ELEMENT WITH SURVEYS

We have argued that a mere showing of likelihood of confusion arising from similarity between the plaintiff's and defendant's trademarks should not be sufficient to establish liability, even if plaintiff's trademark is valid.[104] Rather, to be actionable, confusion must be traceable to *the protected elements* of the plaintiff's trademark. But how can one decide whether confusion is attributable to protected or unprotected elements? How can one trace the source of confusion?

Some evidence of likelihood of confusion does not permit tracing. For example, it is well-established that evidence of *actual* consumer confusion, such as consumers asking Fuddruckers and its competitor whether they were affiliated,[105] is highly probative of likelihood of confusion.[106] And that makes sense: Nothing shows that consumers *would be* confused better than that they *were* confused.[107] The difficulty, however, is that evidence of actual confusion rarely comes with evidence as to the *source* of that confusion. If actual-confusion

102. *See* 1 MCCARTHY, *supra* note 5, § 8:3 ("All courts agree that the elements of the alleged trade dress must be clearly listed and described.") (citing many cases).

103. This is a real trademark. *E.g.*, SEATTLE'S BEST COFFEE, Registration No. 4,629,089.

104. To avoid excessive verbiage, we use "plaintiff" throughout to denote a trademark owner who is complaining that another person, the "defendant," is infringing its trademark.

105. Fuddruckers, Inc. v. Doc's B.R. Others, Inc., 826 F.2d 837, 845 (9th Cir. 1987).

106. *E.g.*, AMF Inc. v. Sleekcraft Boats, 599 F.2d 341, 352 (9th Cir. 1979); Lone Star Steakhouse & Saloon, Inc. v. Alpha of Va., Inc., 43 F.3d 922, 937 (4th Cir. 1995); Morningside Grp. Ltd. v. Morningside Cap. Grp., L.L.C., 182 F.3d 133, 141 (2d Cir. 1999); Beacon Mut. Ins. Co. v. OneBeacon Ins. Grp., 376 F.3d 8, 18 (1st Cir. 2004); Groeneveld Transp. Efficiency, Inc. v. Lubecore Int'l, Inc., 730 F.3d 494, 517 (6th Cir. 2013).

107. *Beacon*, 376 F.3d at 18; *Groeneveld*, 730 F.3d at 517.

evidence reveals nothing about source, and if the confusion could be due to common protectable elements (actionable) or common unprotectable elements (not actionable), this evidence is not as probative as it might appear.

One form of evidence that can tell us not just about the fact of confusion but also about its source is consumer surveys, one of the most important types of evidence in trademark cases.[108] In this Part, we show how to determine the source of confusion with the aid of surveys. We begin with a review of likelihood of confusion surveys and current best practices, then we set forth what we believe constitutes the proper survey methodology and explain how it improves on current wisdom.

A. CONFUSION SURVEYS

In the trademark context, a survey refers to the practice of a researcher or a party's hired expert presenting examples of trademarks or trademarked products to a sample of consumers to elicit their reactions.[109] Surveys can shed light on a variety of questions in trademark litigation, including whether a mark has become generic, whether a mark has secondary meaning, and, most importantly for our purposes, likelihood of confusion.[110] Though rarely used in trademark litigation in the first half of the twentieth century,[111] surveys have become common in recent decades.[112] Reflecting this increased usage, the leading trademark treatise now devotes an entire part of a chapter to surveys,[113] and there is substantial academic and practice-oriented literature on survey

108. *See* Diamond & Franklyn, *supra* note 11, at 2029 ("A consumer survey that measures consumer confusion is an effective way to ensure that trademark infringement cases are decided based on empirical facts about likely consumer confusion instead of on judicial assumptions about how consumers are likely to respond."); 5 MCCARTHY, *supra* note 5, § 32:158 (explaining that surveys provide evidence of public perception and avert the "danger . . . that lawyers and judges will decide these issues of perception by consulting their own personal viewpoint").

109. *See generally* 5 MCCARTHY, *supra* note 5, § 32:158 (providing a primer on trademark surveys).

110. *See* Diamond & Franklyn, *supra* note 11, at 2032–48 (surveying the multiple uses of surveys in trademark and false-advertising litigation); Jerre B. Swann, *Likelihood of Confusion*, *in* TRADEMARK AND DECEPTIVE ADVERTISING SURVEYS: LAW, SCIENCE, AND DESIGN 53, 53 (Shari Seidman Diamond & Jerre B. Swann eds., 2012) ("The plurality of reported Lanham Act surveys address likelihood of confusion.").

111. For example, Judge Jerome Frank, sitting on an all-male panel drawn from an all-male court, lamented the absence of survey evidence from the relevant consumer population in a dispute between the publisher of the "girls' magazine" *Seventeen* and a maker of women's girdles using the trademark "Miss Seventeen." Triangle Publ'ns v. Rohrlich, 167 F.2d 969, 976 (2d Cir. 1948) (Frank, J., dissenting) ("As neither the trial judge nor any member of this court is (or resembles) a teen-age girl or the mother or sister of such a girl, our judicial notice apparatus will not work well unless we feed it with information directly obtained from 'teen-agers' or from their female relatives accustomed to shop for them.").

112. *See* DeRosia, *supra* note 11, at 617 n.7 (reviewing a variety of estimates and concluding that "in trademark litigation, surveys are commonplace but not universal").

113. *See* 5 MCCARTHY, *supra* note 5, §§ 32:158–32:196.

design.[114] The jury is still out on precisely how frequently surveys affect the outcome of trademark cases, but courts and commentators agree that a properly designed survey can be instrumental in proving likelihood of confusion.[115]

Early surveys asked about the plaintiff's and defendant's trademarks or products, inferring likelihood of confusion if a substantial portion of respondents were confused into thinking that the products were affiliated.[116] As for pinpointing the *source* of confusion, early surveys either did not bother[117] or simply asked the respondents who were confused why they were confused.[118]

But the field soon recognized that this practice is not good enough. Asking confused respondents *why* they were confused is not particularly useful because they may not be able to articulate why they thought two products were affiliated. In addition, they often give unhelpful responses such as "similar packaging," which does not make a distinction between protectable and

---

114. *See generally, e.g.*, TRADEMARK AND DECEPTIVE ADVERTISING SURVEYS: LAW, SCIENCE, AND DESIGN, *supra* note 110; Shari Seidman Diamond, *Reference Guide on Survey Research*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 359 (3d ed. 2011); DeRosia, *supra* note 11; Eugene P. Ericksen & Melissa A. Pittaoulis, *Control Groups in Lanham Act Surveys*, 104 TRADEMARK REP. 744 (2014); Jerre B. Swann, *Likelihood of Confusion Studies and the Straitened Scope of Squirt*, 98 TRADEMARK REP. 739 (2008); Jacob Jacoby, *Experimental Design and the Selection of Controls in Trademark and Deceptive Advertising Surveys*, 92 TRADEMARK REP. 890 (2002); Michael Rappeport, *Litigation Surveys—Social "Science" as Evidence*, 92 TRADEMARK REP. 957 (2002).

115. In the 1990s and early 2000s, several commentators, some of whom were frequently hired as survey experts, opined anecdotally that surveys are essential in trademark litigation. *See, e.g.*, Rappeport, *supra* note 114, at 957 (describing surveys as a "necessity"); *Committee Print to Amend the Federal Trademark Dilution Act: Hearing Before the Subcomm. on Cts., the Internet, and Intell. Prop. of the H. Comm. on the Judiciary*, 108th Cong. 14 (2004) (statement of Robert W. Sacoff, Chair, Section of Intell., Prop. L., Am. Bar Ass'n) (characterizing surveys as "one of the most classic and most persuasive and most informative forms of trial evidence"); Itamar Simonson, *The Effect of Survey Method on Likelihood of Confusion Estimates: Conceptual Analysis and Empirical Test*, 83 TRADEMARK REP. 364, 364 (1993) ("Indeed, surveys are now routinely employed to prove likelihood of confusion, and a failure to introduce a survey into evidence often leads to harsh criticism by the courts."). This assessment was challenged by studies based on reported judicial decisions. *See* Beebe, *supra* note 20, at 1622, 1641 (finding that survey evidence was discussed in 20% of the 331 reported federal district court opinions from 2000 to 2004 that used the multifactor test for likelihood of confusion, and that courts rarely credited the proffered survey evidence, and concluding that survey evidence "is in practice of little importance"); 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, 1048 (2012) (concluding, based on 533 federal district court trademark decisions from 2000 to 2006, that "[w]hile surveys can prove persuasive under certain conditions, they might not be useful for litigants with particularly weak or strong evidence"). But looking at reported decisions may not fully capture the impact of surveys because surveys may play an important role in facilitating settlement during the discovery stage. A study considering this possibility found, based mainly on a survey of the members of the International Trademark Association, that surveys are more important than the literature looking only at reported opinions had suggested, especially in closely contested cases. Diamond & Franklyn, *supra* note 11, at 2067.

116. *See, e.g.*, SquirtCo v. Seven-Up Co., 628 F.2d 1086, 1089–90 (8th Cir. 1980); Union Carbide Corp. v. Ever-Ready Inc., 531 F.2d 366, 384–86 (7th Cir. 1976).

117. *See, e.g.*, *SquirtCo*, 628 F.2d at 1089.

118. *Id.* at 1089 n.4; *Ever-Ready*, 531 F.2d at 385.

unprotectable elements of the package.[119] So, something more is needed to get at the source of confusion. That something more is a *control*.[120]

The purpose of a control is to detect and weed out confusion that is due to something other than defendant's use of plaintiff's protected trademark. For example, if consumers think that the parties' products are affiliated just because they are the same product—say, energy bars—such confusion is clearly not actionable. Similarly, as we have argued, even confusion arising from similarity of the parties' trademarks is not actionable if the similarity goes only to an unprotected element of the trademark—for example, the fact that both companies have "bar" in their word mark or both use see-through wrapping.

The way to deal with this problem is to compare the plaintiff's product or trademark not only to the defendant's, which we might think of in this context as the "treatment," but also to another trademark or product that functions as a "control." The literature's current best understanding is that a properly constructed control should come as close as possible to defendant's trademark without containing the elements that plaintiff claims are infringing. Top academic survey experts,[121] the leading trademark treatise,[122] and experts retained for high-profile cases all agree on this approach (although, of course, the last group often disagrees about whether the other side properly followed the agreed-upon approach).[123] The idea is that the treatment (the defendant's mark) yields the "raw" rate of confusion—that is, *all* confusion resulting from defendant's mark, regardless of source. From this raw rate, we subtract the confusion produced by the control, sometimes referred to as "noise"—which, when the control is properly constructed, is nonactionable confusion based on unprotectable similarities. The result is the "net" rate of confusion.[124] This

119. For example, consider some of the responses to plaintiff's expert's survey in the *Kind v. Clif* litigation: "Same or similar packaging as the earlier snack bars." "Very similar packaging." "They look the same." "The same package." "The package is almost identical." "The packaging of the two brands of snack bars are extremely similar." "Looks like the same kind packaging." "Looks like the same type wrapper." *See* Declaration of George Mantis ¶ 15, Kind LLC v. Clif Bar & Co., No. 14 Civ. 770 (S.D.N.Y. Feb. 12, 2014), ECF No. 14, Exhibit F [hereinafter Mantis Declaration].

120. *See, e.g.*, Shari Seidman Diamond, *Control Foundations: Rationales and Approaches*, *in* TRADEMARK AND DECEPTIVE ADVERTISING SURVEYS: LAW, SCIENCE, AND DESIGN, *supra* note 110, at 201, 201; 5 MCCARTHY, *supra* note 5, § 32:187.

121. Diamond, *supra* note 120, at 210 ("The general principle for choosing an appropriate control is easily stated: It should share as many characteristics with the experimental stimulus as possible, with the key exception of the characteristic whose influence is being assessed.").

122. 5 MCCARTHY, *supra* note 5, § 32:187.

123. This was true of the experts on both sides of *Kind v. Clif*, who agreed on the principle of control selection but disagreed as to the appropriateness of the control selected by the plaintiff's expert. *See* Mantis Declaration, *supra* note 119, ¶ 11 ("The best control is one that holds constant every variable that may affect respondents' answers, except for the variable (trade dress) being tested."); Declaration of Michael Rappeport at 6, Kind LLC v. Clif Bar & Co., No. 14 Civ. 770 (S.D.N.Y. Mar. 31, 2014), ECF No. 35 [hereinafter Rappeport Declaration] ("[A] 'best' control is that control which comes as close as possible to the claimed infringing usage without itself being infringing.").

124. 5 MCCARTHY, *supra* note 5, § 32:187.

net rate then measures that part of the confusion that is attributable to defendant's allegedly infringing conduct, which is the relevant outcome from the survey.[125]

To see how this works, let us go back to the energy bar example. Suppose Kind, which makes an energy bar with wrapping p, is suing Clif, which makes an energy bar with wrapping d, for trade dress infringement. (All of our notation in this Article is collected in a Notation Table in Section C of the Appendix.) Suppose further that p *both* has an overall configuration of elements—a layout of shapes and colors—that is trademarkable *and* has some see-through wrapping, which by itself is not trademarkable. Now if d also contains both of these features or something similar to them, a survey that showed respondents only d (or only p and d) and found that a substantial portion were confused into thinking that p and d are from an affiliated source would not be able to tell whether the confusion is due to the similarity of the overall configuration (actionable) or to the fact that both used see-through wrapping (not actionable). So, to adjudicate between these competing explanations, the analyst would introduce a control energy bar, let's call it d^u, that has the (unprotectable) see-through wrapping but not the (protectable) overall configuration. The analyst would then subtract the "noise" confusion rate given by the control experiment from the "raw" confusion rate in the initial experiment to arrive at the "net" confusion rate that captures the true quantity of interest—confusion attributable to the protectable overall configuration. For example, if 20% of respondents were confused in the initial experiment involving d and 10% were confused in the control experiment involving d^u then the net confusion rate would be 10%. That is the approach currently advocated by top survey experts and embraced by the leading trademark treatise.[126]

Some (though by no means all) courts have gotten fairly sophisticated about the design of trademark confusion surveys, scrutinizing the parties' choice of control to make sure it enables the proper comparison.[127] Courts are helped

125. An unstated (though eminently reasonable) assumption behind this approach is that if the control is properly constructed, the raw rate of confusion is greater than or equal to the noise, yielding a nonnegative net rate of confusion. That is, if a consumer is not confused by the defendant's product, they would also not be confused by a control product that includes *fewer* elements in common with the plaintiff's trademark. More generally, this assumption, which we call "monotonicity," is as follows: With more similarities, confusion will only stay the same or increase; it will not decrease. We retain this assumption for purposes of our own proposed methodology in Section II.B. Without it, the use of controls altogether would be questionable.

126. *See* 5 MCCARTHY, *supra* note 5, § 32:187; *see also, e.g.*, Mantis Declaration, *supra* note 119, ¶¶ 15–17 (subtracting the control group's confusion rate from the treatment group's confusion rate to arrive at the net confusion rate that is reported as the relevant rate); Rappeport Declaration, *supra* note 123, at 10 (following the same approach).

127. *See, e.g.*, Valador, Inc. v. HTC Corp., 242 F. Supp. 3d 448, 463–64 (E.D. Va. 2017), *aff'd*, 707 F. App'x 138 (4th Cir. 2017) (excluding an expert because, among other things, his survey did not include a control); U.S. Polo Ass'n v. PRL USA Holdings, Inc., 800 F. Supp. 2d 515, 534–35 (S.D.N.Y. 2011), *aff'd*, 511 F. App'x 81 (2d Cir. 2013) (giving a survey "no weight" in part because

in this regard by the parties' survey experts, who critique the other side's survey methodology. The battle of the experts in the *Kind v. Clif* case is instructive. In that case, Kind claimed that its trade dress (Figure 3) was infringed by Clif's Mojo Bar (Figure 4). Kind's expert conducted a survey experiment with a control, finding a raw confusion rate of 26% and a control confusion rate of 11%, amounting to a net confusion rate of 15%.[128] The control used by Kind's expert showed an old design of the Clif Mojo Bar to which Kind had taken no objection (Figure 5). Kind's expert claimed to follow the above-described approach to control selection: "The best control is one that holds constant every variable that may affect respondents' answers, except for the variable (trade dress) being tested."[129] He explained that "the CLIF control bar [Figure 5] is an ideal control because, as the same type of product sold under the same brand name but with a completely different trade dress, it holds constant every variable except trade dress."[130] Clif's expert took the same approach to control selection[131] but objected to his adversary's chosen control.[132]

The problem with Kind's expert's control, explained Clif's expert, is that it did not contain *unprotectable elements* of the Kind Bar trade dress, most notably the see-through wrapping; as such, it could not filter out confusion that was due to the similarity of unprotectable elements and isolate the confusion attributable to protectable elements.[133] To correct this error, Clif's expert chose controls that not only were a noninfringing energy bar (like Kind's expert's control) but also contained the see-through wrapping (Figure 6).[134] Clif's expert thus arrived at a larger estimate of confusion noise (26% or

its "controls were improper in that they included the very elements being assessed"); 24 Hour Fitness USA, Inc. v. 24/7 Tribeca Fitness, LLC., 447 F. Supp. 2d 266, 280 (S.D.N.Y. 2006), *aff'd*, 247 F. App'x 232 (2d Cir. 2007) (finding "compelling" defendant's expert's argument that plaintiff's control, "Lifetime Fitness," was inadequate, because it did not express that the gym was open 24 hours a day). Courts often decline to exclude under Federal Rule of Evidence 702 surveys with improper controls, reasoning that the flaws can be uncovered through cross-examination. *See* Saxon Glass Techs., Inc. v. Apple Inc., 393 F. Supp. 3d 270, 290–91 (W.D.N.Y. 2019), *aff'd*, 824 F. App'x 75 (2d Cir. 2020). Our aim here is to pinpoint and remedy a flaw in survey methodology; we do not take a position on whether and when a survey that deviates from our methodology is not "reliable" and should be excluded under Rule 702.

128. Mantis Declaration, *supra* note 119, ¶¶ 15–17.

129. *Id.* ¶ 11.

130. *Id.*

131. *See* Rappeport Declaration, *supra* note 123, at 6 ("[A] 'best' control is that control which comes as close as possible to the claimed infringing usage without itself being infringing.").

132. *Id.* at 2 (explaining that his survey "replicates the basic . . . survey approach" used by Kind's expert, except for the choice of controls); *accord id.* at 5–6.

133. *Id.* at 5.

134. *Id.* at 6–7. Clif's expert actually used two controls, but we show only one of them here so as not to clutter this Section with pictures (more pictures appear in our survey instrument in Part III). Although two different controls were used, each control-group respondent was shown only one of the controls. *Id.* at 6 n.9.

21%, depending on the control), resulting in a reduced rate of net confusion (0 or 5%).[135]

The *Kind v. Clif* example shows real progress in likelihood of confusion surveys. The field has come far from the days of no surveys to surveys without controls to surveys with increasingly sophisticated controls. It might even seem, in light of Clif's expert's critique, that at least some sophisticated practitioners are already filling in the "missing element" we have identified by carefully

Figure 3: Kind's Allegedly Infringed Trade Dress

Figure 4: Clif's Allegedly Infringing Trade Dress

Figure 5: Kind's Control

Figure 6: One of Clif's Controls

selecting controls. Unfortunately, that is not so. As we explain next, the best existing approaches to the design of trademark surveys, including both parties' approaches in the *Kind v. Clif* litigation, are still inadequate.

135. *Id.* at 10.

B. A Better Approach to Confusion Surveys

We argue that existing approaches are not well-designed to tease out different possible sources of confusion and, therefore, do not permit a well-identified estimate of actionable confusion. As we shall explain, a well-identified survey requires not one control but two or more controls. Before turning to the nub of our methodological contribution, two preliminary points are in order.

First, we use a *between-subject* survey design, where each respondent is exposed to either the defendant's trademark or one of the controls, as opposed to a *within-subject* design, where the same respondents are exposed to both the defendant's product and the controls.[136] This allows us to ensure that the controls do not influence how respondents react to the defendant's product (or vice versa),[137] and it allows us to cleanly isolate the confusion rates for individual controls.[138] Although both between-subject and within-subject survey designs are used in trademark litigation,[139] our between-subject design best allows us to differentiate between possible sources of confusion. If we *were* able to craft a reliable within-subject survey, however, our method would work for that design, as well.[140]

Second, properly designing a survey involves a host of additional considerations—the selection of respondents, the wording of questions, and many other things—but we abstract away from those in this Section to focus on the design of controls, which is the nub of our methodological contribution. We are assuming, in effect, that the surveys are properly designed in all other

136. *See* Gary Charness, Uri Gneezy & Michael A. Kuhn, *Experimental Methods: Between-Subject and Within-Subject Design*, 81 J. ECON. BEHAVIOR & ORG. 1, 1 (2012) (defining "within-subject" and "between-subject" experiment designs).

137. Between-subject designs "are the strongest designs for causal inference because a respondent cannot be influenced by the content of material shown to respondents in other cells of the design." Diamond, *supra* note 120, at 216. By contrast, a respondent *can* be influenced by the content of material that they have already seen or that they view simultaneously with the product of interest, as they would in a within-subject survey. In other words, respondents' responses when faced with a control or treatment condition are not independent of having also been exposed to other conditions. This risk is particularly high in our experimental setting, where the defendant's product and one of the controls have very similar trade dresses.

138. *See* Charness et al., *supra* note 136, at 2 (noting that exposing each subject to multiple treatments may introduce "a slew of confounds to identification").

139. *See* Mike Rappeport, *Design Issues for Controls*, *in* TRADEMARK AND DECEPTIVE ADVERTISING SURVEYS: LAW, SCIENCE, AND DESIGN, *supra* note 110, at 217, 236–39 (describing "array" surveys where, for example, respondents are first shown the plaintiff's product, then shown an array that includes both the defendant's product and controls, then asked questions to determine whether they believe any of the products came from the same source as the plaintiff's product); Jacoby, *supra* note 114, at 912 (observing that "within-group designs relying upon internal controls are often employed in trademark surveys," including array surveys). These arrays are used to minimize "demand effects." *See infra* text accompanying notes 191–97. In our sample survey, we use a different method—decoy product types—to minimize demand effects. *See infra* text accompanying notes 200–02.

140. *See infra* note 220. We thank Scott Baker and Jens Frankenreiter for suggesting a discussion of between-subject and within-subject designs.

respects. The other considerations do, however, enter into the design of our survey in Part III.[141]

1. A Simplified World: No Confusion from Inevitable Product Features

We begin by considering a context where any confusion as to the source or affiliation of the products in litigation must come from their trademarks or from features of the product that might vary across different producers. That is, consumer confusion due to *inevitable non-trademark* features of the products (or due to just being naturally confused) is nonexistent or negligible; no consumer thinks that the plaintiff's and defendant's products come from the same source just because they are the same product type or just because they share some feature that is inherent to the product type. There are some contexts where this assumption seems reasonable. For example, if one car manufacturer sues another for using a confusingly similar trademark, there will be negligibly few consumers who make an inference of affiliation purely on the basis that both products are cars or that both have wheels. In other (perhaps most) contexts, this assumption is not a safe one, and those other contexts are considered below in Section II.B.2. But we begin with the simpler context because it more clearly illuminates the proper approach and the shortcomings of existing approaches. Note that, even in the simplified context, confusion due to *non-protectable elements* of the plaintiff's trademark (or even non-trademark but not inevitable features) is possible; the only thing we rule out by assumption is confusion due to *inevitable non-trademark* features. So, in the energy bar example, people may be confused by, say, the transparent packaging but not by the mere fact that both plaintiff and defendant make energy bars.

In systematically laying out the proper approach to survey design, it is helpful to begin with an abstract formulation of the problem at hand. This does not involve any advanced mathematics, but it requires some notation to enhance clarity and precision.

The problem can be formulated as follows. A senior user P, who makes a product with trade dress p, is suing a junior user D, who makes a product with trade dress d, for trademark infringement. Generally speaking, one might conceive of p as an n-tuple $(p_1, \ldots, p_n)$ and of d as an m-tuple $(d_1, \ldots, d_m)$. For ease of exposition, and without loss of generality, consider the case where both trade dresses have three elements, so we have (p_1, p_2, p_3) and (d_1, d_2, d_3), and the trade dresses are similar with respect to the first two elements, so d_1

141. We do have one serious reservation about trademark confusion surveys, and that is their use—or rather non-use—of statistics. Survey experts interpret their experimental results without any reference to statistical significance, in effect treating parameter estimates from their sample as the true population parameter. The survey literature as of yet does not seem to have given serious thought to questions of statistical inference. This is a deeper problem than we can tackle here, but we hope to explore it in the future.

is like $p_1$and d_2 is like p_2.[142] Suppose further that element p_1 is by itself unprotectable (e.g., see-through packaging for energy bars) but p_2 is protectable (e.g., the overall configuration of the packaging). As for p_3, it can stand for any elements in the plaintiff's trade dress, protectable or unprotectable, that are not similar to defendant's. Because p_1 is *unprotectable*, we will rename it p_u, and because p_2 is *protectable*, we rename it p_p. So, the plaintiff's trade dress can be expressed as (p_u, p_p, p_3). To make this concrete, we will continue to run with the energy bar example and assume that p_u is see-through wrapping (not protectable), p_p is the overall configuration (protectable), and p_3 covers the other elements of the Kind trade dress (including, say, the Kind logo) that are not similar to the Clif trade dress.

Note that, as defined, the "elements" of the allegedly infringed trade dress do not have an *a priori* or Platonic definition; rather, the elements are defined in the context of a given case, a given allegedly infringing trade dress.[143] Also note that a distinctive overall configuration of the trade dress could itself be one of the elements, as is the case in our leading example. Finally, note that the approach assumes clarity on what the protectable and unprotectable elements are in the plaintiff's trade dress. This determination is part of trademark *validity* analysis,[144] not likelihood of confusion, so all approaches to survey design assume some basic agreement on what is or is not protectable. That is not to say that pinpointing what is protectable about a mark is a trivial task; for some marks, it may be quite difficult to precisely delineate the scope of protection. It would behoove the parties and the court to achieve a common understanding of this scope before their experts conduct surveys.[145] But whether the parties agree or not, scope delineation is a prerequisite to any analysis tracing the source of confusion—whether that analysis uses our proposed survey method, some other survey method, or any non-survey method for that matter.[146]

---

142. The reason moving from the general *n*-versus-*m*-dimensional case to the specific three-dimensional case does not lose generality is that, for present purposes, we can collapse some of the elements into one. For example, we can collapse into p_1 all unprotectable trademark elements as well as all non-trademark, variable features of the plaintiff's product that are in common with the defendant's.

143. Any combination producing an overall look and feel can be divided into constituent elements in any number of different ways, none intrinsically superior to another, so an analysis that required a particular subdivision would not be stable. Our approach avoids that problem by defining the elements in a functional way that corresponds to the particular case at hand. So, for example, it is theoretically possible that the same visual feature would count as element p_p in one case and p_3 in another case.

144. *See supra* notes 93–101 and accompanying text (discussing scope determination).

145. *See* Diamond, *supra* note 120, at 215 (noting that survey experts should be "as clear as possible" about the scope of trademark protection "before designing a survey" and that judges can help the parties achieve clarity on the issue). A "scope proceeding" like the one suggested by Lemley and McKenna, see *supra* note 21, may be helpful to this task. *See* text accompanying notes 101–02.

146. If prior clarity on the scope of trademark protection cannot be achieved by agreement of the parties or the court's order, the parties may each do their own survey based on *their* understanding

Let us now work through how one should approach a trademark infringement claim in the context of different possible configurations of d, keeping in mind the requirement that, to be actionable, confusion must be attributable to protectable elements.

Case 1: Defendant's product contains only an unprotectable element of plaintiff's trademark. Suppose first that the only thing d shares with p is the unprotectable element (in the energy bar example, see-through wrapping). That is, d takes the form (p_u, d_2, d_3). (When we write an element of d as an element of p—that is, when we write d_i as p_i—we mean that in this (ith) element d and p are the same or very similar.) In this case there can be no actionable confusion, and we don't even need a survey to know that. For even if a substantial portion of consumers are likely to be confused into thinking that p and d are affiliated, that confusion can only be attributed to an unprotectable element, which, as explained in Part I, cannot give rise to trademark infringement. The reason we can pin the confusion down to the unprotectable element is that that's the only thing p and d have in common.

Case 2: Defendant's product contains only a protectable element of plaintiff's trademark. Suppose next that d shares with p only a protectable element of the trade dress. That is, d takes the form (d_1, p_p, d_3), which in the energy bar example translates to sharing the overall configuration of the packaging. In this case, a survey would be useful in gauging whether an appreciable portion or number of consumers would be confused into thinking that p and d are affiliated in source or sponsorship.[147] And all it takes is a survey involving d; we do not need a control. If the survey involving d shows that an appreciable portion of consumers think that p and d are affiliated, then it enables an inference of likelihood of confusion, and if it does not show that, then it does not enable such an inference.[148] The reason we do not need a control is that any confusion between p and d can be attributed to the protectable element p_p because that is the only thing p and d have in common (except for inevitable non-trademark features, which we have assumed cannot cause any non-negligible confusion). So, the confusion, if it exists, is indeed actionable.[149]

---

of what is and is not protectable. The court would then have to adjudicate both the scope of protection and the merits of the surveys. Whatever the proper scope of protection, we argue that our method is the best way to trace how much confusion is attributable to the protectable parts of the trademark.

147. What constitutes "appreciable" depends on the case, including how the eight-factor likelihood-of-confusion inquiry has shaken out. *See* 5 MCCARTHY, *supra* note 5, §§ 32:185, 32:188. We take the necessary threshold of confusion as given; our approach would work given any threshold.

148. We are assuming, as we said, that the survey is otherwise properly designed.

149. Our assumption that no non-negligible confusion can result from inevitable non-trademark features is doing real work here. For although p_p is the only *trademark* thing p and d have in common, they may have other, *non*-trademark things in common (say, the fact that both are rectangular energy bars). If those other things could cause non-negligible consumer confusion—a possibility we have ruled out by assumption in this Section—then a control would be needed to tease out the source of confusion in this case. We address the proper design of such a control in Section II.B.2.

Case 3: Defendant's product contains both unprotectable and protectable elements of plaintiff's trademark. Finally, assume that d shares both unprotectable and protectable elements with p. That is, d takes the form (p_u, p_p, d_3), which in the energy bar example translates to the defendant's packaging having similar elements to the plaintiff's unprotectable see-through wrapping and protectable overall configuration, but not the protectable Kind logo. This is the most interesting problem, and it's where we make our main contribution to survey methodology.

We begin with a survey involving d. If the survey does not show that an appreciable portion of respondents are confused into thinking that p and d are affiliated, then likelihood of confusion cannot be inferred, and no controls or further surveys are needed. But what if the survey shows that an appreciable portion of consumers are confused? What more do we need to pin down the source of confusion and decide whether it might be actionable?

As we saw in Section II.A, the most advanced version of current wisdom is that we need a control that comes as close as possible to d without containing the protectable element on which d is allegedly stepping. In our notation, we need a control, call it d^u, of the form (p_u, d_2, d_3). This translates in the energy bar example to a control that has the see-through packaging but not the overall configuration (nor the logo). We subtract the confusion rate between p and d^u (so-called "noise") from the confusion rate between p and d ("raw" confusion). If the result ("net" confusion) still leaves an appreciable portion of confused respondents, we can infer likelihood of confusion; otherwise, we infer no likelihood of confusion.[150]

We claim that the second deduction does not logically follow. That is, although we agree that likelihood of confusion can be inferred if subtracting the p/d^u confusion rate[151] from the p/d confusion rate leaves an appreciable net rate of confusion, we do not agree that likelihood of confusion should be *ruled out* if subtracting the p/d^u confusion rate from the p/d confusion rate does *not* leave an appreciable net rate of confusion. That is so because the fact that the substraction does not leave appreciable confusion does not, on its own, rule out the possibility of appreciable confusion due to p_p alone—which *would* be actionable. And if there is appreciable confusion due to p_p alone, then that would render the appreciable confusion in the initial p/d survey actionable, even if there is also nonactionable confusion due to p_u alone.

In other words, if a respondent is confused both by the inclusion of p_u (the see-through wrapping) alone and by the inclusion of p_p (the overall configuration) alone, then that respondent should be included in the portion of those who are actionably confused when d includes both p_u and p_p. A respondent should count as actionably confused if p_p is *sufficient* to cause her

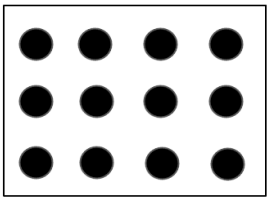
150. *See supra* notes 119–25 and accompanying text.

151. We use the notation p/d as shorthand for the survey measuring confusion between p and d, so "the p/d confusion rate" means the rate of confusion measured in such a survey (and similarly for p/d^u and so on).

confusion, even if p_u is *also* sufficient to cause her confusion.[152] But the current approach would subtract that respondent from the group of those actionably confused, and therein lies its error.[153] To rephrase (again), in assessing the confusion resulting from the inclusion of both p_u and p_p, the existence of nonactionable confusion due solely to p_u, the unprotectable element, does not render any confusion due solely to p_p, the common protectable element, nonactionable. Rather, the latter confusion, if it exists, is actionable and would render the confusion resulting from the inclusion of both p_u and p_p actionable.

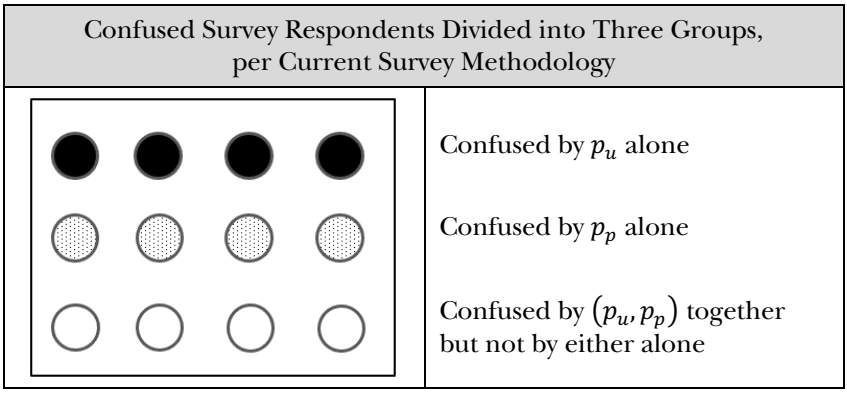
A visual representation illustrates the point. Think of the group of survey respondents who are confused between the plaintiff's and defendant's products as represented by these balls:

Figure 7: Confused Survey Respondents, Represented by Balls



The current approach divides the balls into three categories, each represented by a different shading below:

Figure 8



First, the solid black balls in Figure 8 represent people who are confused by the unprotectable features standing alone. In our example, these are people who would think that the defendant's product was affiliated with the plaintiff's if the only similarity were the (unprotectable) see-through wrapping. Second,

152. We defend this view of factual causation in Section II.C.

153. See the Appendix for a formal demonstration of how the current approach goes wrong.

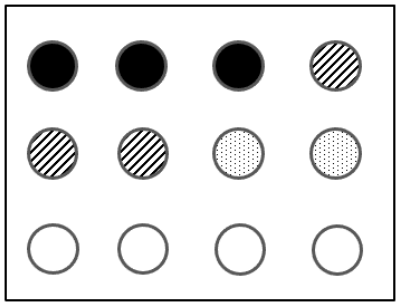
the dotted balls represent people who are confused by the protectable features standing alone. In the example, these are people who would think that the defendant's product was affiliated with the plaintiff's if the only similarity were the (protectable) overall configuration. Third, the white balls represent people who are confused by the protectable and unprotectable features together but not by either alone. These people would be confused into thinking that the defendant's energy bar was affiliated with the plaintiff's only if it had both (unprotectable) see-through wrapping and a similar (protectable) overall configuration.

The current best practice is to isolate and subtract the black balls from the rate of confusion. This can be accomplished with the d^u control: People confused by d^u are confused by the unprotectable features alone—they are the solid black balls. Some experimental subjects will see d ("condition d") and some will see d^u ("condition d^u"). By subtracting the rate of confusion in condition d^u from the rate of confusion in condition d, the expert subtracts out the black balls. The idea behind current best practices is that by taking out the black balls, we take out the "noise"—the people confused by the unprotectable features—leaving only the "signal," the people confused by the protectable features.

Our point is that this approach is wrong. Specifically, it is wrong because the set of people who are confused by the unprotectable features standing alone (the solid black balls) and the set of people who are confused by the protectable elements standing alone (the dotted balls) may overlap. There might be people who would be confused *both* by the unprotectable features standing alone *and* by the protectable features standing alone. In our example, a single person might think that an energy bar is affiliated with the plaintiff if it has *either* (unprotectable) see-through wrapping *or* a similar (protectable) overall configuration. This person meets both the black-ball criterion and the dotted-ball criterion, meaning they cannot be sorted into a single set.[154] Therefore, instead of the three-part division above, the group of respondents confused by defendant's product should be divided into four groups, like so:

154. In other words, the current best practice's attempted partition of the set of confused people is not a true partition because the subsets are not mutually disjoint. See the Appendix for a more detailed discussion.

Figure 9

Confused Survey Respondents Divided into Four Groups, per Our Survey Methodology	
	 Confused by p_u alone but not by p_p alone
	 Confused both by p_u alone and by p_p alone
	 Confused by p_p alone but not by p_u alone
	 Confused by (p_u, p_p) together but not by either alone

This new partition adds the striped balls: people who would be confused both by the protectable features alone and by the unprotectable features alone. It also modifies the definition of the black balls: Those are now people who are confused by the unprotectable elements alone *but not by the protectable elements alone*. The definition of the dotted balls is similarly modified: They are people confused by the protectable elements alone but not by the unprotectable elements alone. This is now a true partition: Every person confused between the plaintiff's product and the defendant's product can be sorted into one and only one set.[155] The striped-ball people should count toward the actionable rate of confusion, because their confusion is traceable to the inclusion of p_p, the protectable features.[156] Therefore, the true actionable rate of confusion is the striped balls plus the dotted balls plus the white balls; it consists of everyone confused by p_p alone and everyone confused by (p_u, p_p) together but not by either alone.

But how do we know if there is appreciable confusion due to p_p alone? How do we separate the striped-ball people from the solid-black-ball people? The answer is: We need another control. Take a survey involving a second control, d^p, of the form (d_1, p_p, d_3). In our energy bar example, this means a control that shares the protectable overall configuration but not the unprotectable see-through wrapping (nor the protectable logo or other features) with the plaintiff's trade dress. If the survey shows appreciable confusion between p and d^p, then that finding, combined with the earlier finding of appreciable confusion from the "treatment" survey involving d, justifies an inference of likelihood of confusion, because these findings together show that there is appreciable confusion between p and d that is attributable to a protected element of p. But if the survey does not show appreciable confusion between p and d^p, then that finding, combined with the earlier finding that the difference

155. For a more detailed and formal discussion of this partition, see the Appendix.

156. For a defense of this position, see *infra* Section II.C.

between the p/d and p/d^u confusion rates is not appreciable, implies that we cannot infer actionable likelihood of confusion. That is so because these findings, taken together, show that we cannot attribute appreciable confusion between p and d to a protectable element of p.[157]

Note that, like the first control, the second control by itself (that is, standing alongside the initial treatment survey finding appreciable confusion) cannot be enough to rule out an inference of likelihood of confusion. This is because a finding that there is no appreciable confusion between p and d^p, though ruling out an inference of confusion due to p_p alone, does not rule out the possibility that appreciable confusion is caused by a *combination* of p_u and p_p, which would be actionable. But if we find that there is no appreciable confusion between p and d^p *and* that the difference between the p/d and p/d^u confusion rates is not appreciable, then we can be more confident in rejecting an inference of likelihood of confusion.[158]

The following decision tree summarizes the conclusions from the preceding analysis.[159]

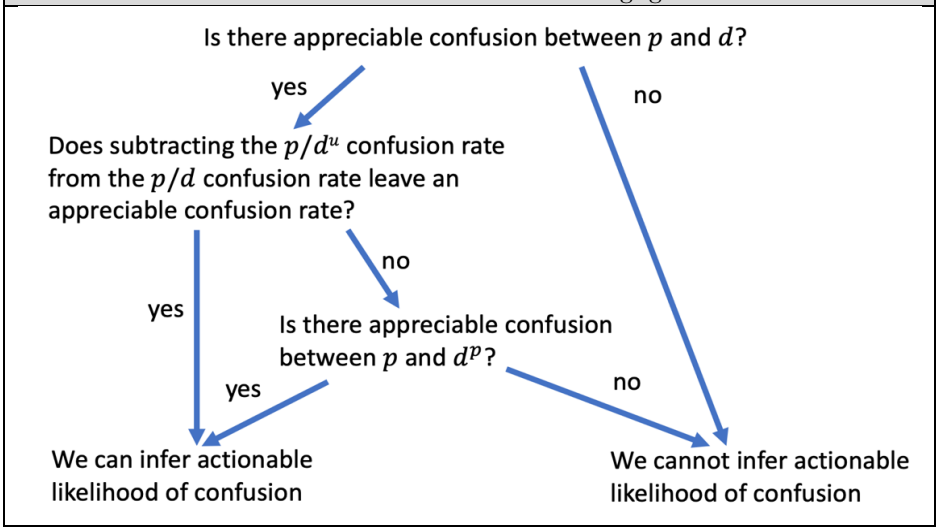
157. Strictly speaking, what these findings show is not that we can *rule out* actionable likelihood of confusion but that we can be more confident in concluding that there is no actionable likelihood of confusion. See the Appendix for a more precise statement and explanation.

158. It should also be clear that the temporal ordering of the control surveys does not matter. We discussed the d^u control first for expositional reasons, since that is the control the current literature has latched onto, but we could have switched the order. One could do either the d^u or the d^p survey first, or one could do them simultaneously with different groups of respondents. The important point is that we can be more confident in *failing* to infer actionable likelihood of confusion when we use both control surveys rather than just one control survey, even though each control survey standing alone (that is, standing alongside the initial treatment survey finding appreciable confusion) can justify an inference of actionable likelihood of confusion.

159. As mentioned, the ordering of the second and third questions can be reversed. *See supra* note 158.

Figure 10

Decision Tree on Whether to Infer Actionable Likelihood of Confusion from Surveys Where Confusion Due to Inevitable Product Features Is Negligible



And the following table shows what conclusions can be reached depending on the outcomes of the control surveys, along with the reason behind the conclusions (column 3). The table assumes that the initial treatment survey has found appreciable confusion between p and d (otherwise, we cannot infer likelihood of confusion regardless of the control surveys' outcomes).

Table 1

Whether to Infer Actionable Likelihood of Confusion Based on Different Outcomes of Control Surveys Where the Treatment Survey Has Shown Appreciable Confusion and Confusion Due to Inevitable Product Features Is Negligible			
Appreciable difference in confusion between p/d and p/d^u?	Appreciable confusion in p/d^p?	Cause of confusion	Result
Yes	Yes	Both p_p alone and (p_u, p_p) together are confusing, and p_u alone may or may not be confusing but in either event is appreciably less confusing than (p_u, p_p) together.	Infer actionable likelihood of confusion
Yes	No	(p_u, p_p) together is confusing, p_p alone is not confusing, and p_u alone may or may not be confusing but in either event is appreciably less confusing than (p_u, p_p) together.[160]	Infer actionable likelihood of confusion
No	Yes	(p_u, p_p) together is confusing but not appreciably more so than p_u alone, p_p alone is confusing, and p_u is also confusing if the confusion in p/d^u is appreciable.[161]	Infer actionable likelihood of confusion
No	No	(p_u, p_p) together is confusing but not appreciably more so than p_u alone, p_p alone is not confusing, and p_u is confusing if the confusion in p/d^u is appreciable.[162]	Cannot infer actionable likelihood of confusion

160. Combining the facts that (1) the p/d confusion rate is appreciable; (2) the difference between the p/d and p/d^u confusion rates is appreciable; and (3) the p/d^p confusion rate is not appreciable does *not* necessarily imply that the p/d^u confusion rate is not appreciable. To see this, suppose the threshold for appreciable confusion is 20%, the p/d confusion rate is 60%, the p/d^u confusion rate is 30%, and the p/d^p confusion rate is 10%. Of course, the p/d^u and p/d^p confusion rates need not add up to the p/d confusion rate. The whole can be more than the sum of its parts; that, in a sense, is the very idea behind protecting overall look and feel trade dress.

161. The fact that the p/d confusion rate is appreciable combined with the fact that there is no appreciable difference between the p/d and p/d^u confusion rates does *not* necessarily imply that the p/d^u confusion rate is appreciable. To see this, suppose that the threshold for appreciable confusion is 20%, the p/d confusion rate is 25%, and the p/d^u confusion rate is 15%.

162. Combining the facts that (1) the p/d confusion rate is appreciable; (2) the difference between the p/d and p/d^u confusion rates is not appreciable; and (3) the p/d^p confusion rate is not appreciable does *not* necessarily imply that the p/d^u confusion rate is appreciable. To see this, suppose the threshold for appreciable confusion is 20%, the p/d confusion rate is 25%, and the p/d^u and p/d^p confusion rates are 10% each.

The analysis so far has shown *whether* actionable likelihood of confusion can be inferred from various possible combinations of results from the treatment and control surveys, but we have not said anything about estimating the *rate* of actionable confusion. In fact, a court may be interested not only in a binary determination of whether or not the threshold of appreciable confusion is met but also in an estimate of how far above or below the threshold the relevant rate falls, so an estimate of the rate would be helpful. We have shown why common estimates of the actionable rate of confusion are wrong—both the naïve estimate of taking the p/d confusion rate (an overestimate) and the field's current best practice of subtracting the p/d^u confusion rate from the p/d confusion rate (an underestimate)—but we have not said what the right estimate is. Deriving the correct estimate turns out to be somewhat subtle. We leave the task to the Appendix, where we also delve into some of the nuances that the foregoing analysis sidestepped.[163] We do, however, incorporate some of the conclusions in the Appendix into our analysis of hypothetical survey results in Part III below.

2. The Real World: Confusion May Come from Inevitable Product Features

We now consider the more realistic context where confusion can also be due to inevitable non-trademark features of the products. The analysis is a bit more involved than the simplified world of Section II.B.1, but the essential insights are very similar. Because the reader is now familiar with the logic of the analysis, we will move more rapidly through the explanations. And because the insights are so similar, readers eager to read about causation may skip to Section II.C.

The context, again, is that senior user P, who makes a product with trade dress p, is suing junior user D, who makes a product with trade dress d, for trademark infringement. This time, confusion may be caused not just by elements of the trade dress or other variable features of the products but also by shared or similar features of p and d that are inherent to the product. For example, some survey respondents might be so confused as to think that any two makers of energy bars are affiliated or that any two products shown in a survey are from affiliated entities ("demand effects")[164]—hence the idea that some potential sources of confusion are "inevitable" in the sense that one cannot design a comparable control product that does *not* contain this potential source.

To capture this possibility, we add an element to the description of p that we call p_0. For example, where the trade dress is the packaging of an energy bar, p_0 captures the product being an energy bar. So p is now represented by

163. Following the argument in the Appendix does not require any knowledge of mathematics or statistics, and we strongly encourage interested readers to read it. But we are also mindful of readers who may not be interested in methodological details, so we chose not to include the rate estimations analysis in the main body of the paper.

164. For a discussion of demand effects, see *infra* text accompanying notes 193–96.

(p_0, p_u, p_p, p_3), and d and all the controls also contain p_0. Let us now work through the same hypothetical scenarios as in the simplified world (Section II.B.1) to see how, if at all, the analysis changes.

Case 1: Defendant's product contains only an unprotectable element of plaintiff's trademark. That is, d takes the form (p_0, p_u, d_2, d_3). The answer is the same as in the simplified world: Because the only commonalities between p and d are unprotectable, there can be no actionable likelihood of confusion. And we don't need a survey to tell us that.

Case 2: Defendant's product contains only a protectable element of plaintiff's trademark.[165] That is, d takes the form (p_0, d_1, p_p, d_3). In the simplified world, where confusion due to non-trademark similarities (the element p_0) was ruled out by assumption, we said that a survey involving d, without any controls, is sufficient. That is no longer true. Because confusion between p and d may now result not only from p_p but also from p_0, the p/d survey may not be enough. Of course, if the p/d survey reveals no appreciable confusion, then the inquiry is at an end, and we can infer no likelihood of confusion. But if the p/d survey does reveal appreciable confusion, then it is no longer sound to infer actionable likelihood of confusion because the confusion may be attributable to p_0.

To tease out the source of confusion, we need a control of the form (p_0, d_1, d_2, d_3), which we call d^0. For example, d^0 could be an energy bar in rectangular package that does not have anything else in common with plaintiff's package. If the difference between the p/d and p/d^0 confusion rates is appreciable, then we can infer actionable likelihood of confusion because p_p is responsible for appreciable confusion between p and d; but if the difference between the p/d and p/d^0 confusion rates is not appreciable, then we cannot infer actionable likelihood of confusion because we cannot determine that p_p's contribution to confusion between p and d is appreciable.

Case 3: Defendant's product contains both unprotectable and protectable elements of plaintiff's trademark. That is, d takes the form (p_0, p_u, p_p, d_3). As in the simplified world, if the treatment survey does not show any appreciable confusion between p and d then no likelihood of confusion can be inferred. But if the treatment survey does reveal appreciable confusion, then here is how we should go about determining whether the confusion is actionable.

We begin the analysis with control d^u of the form (p_0, p_u, d_2, d_3);[166] in other words, this control shares only the unprotectable features (including the inevitable features) of the plaintiff's mark. Note that this is analogous to d^u in the simplified world of Section II.B.1. If subtracting the p/d^u confusion rate from the p/d confusion rate leaves an appreciable rate of confusion, then

165. Readers who have worked through the Appendix will notice that the analysis of Case 2 and Case 3 in this Section elides some nuances (though the elisions do not affect our main point nor make anything we say wrong).

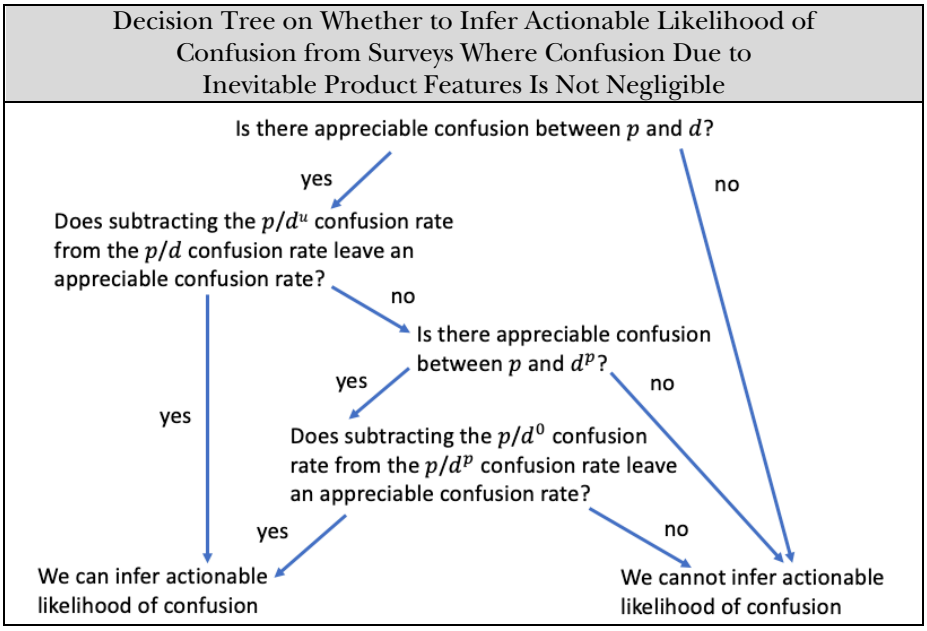
166. Again, this is just how we are structuring the sequence of analysis for expositional purposes; it need not be the temporal sequence of the survey experiment, which might administer the different control conditions in a different sequence or simultaneously to different groups.

actionable likelihood of confusion can be inferred because the protectable element p_p is responsible for appreciable confusion. But, by the same reasoning as in Section II.B.1, and contrary to current best practices and to defendant's expert in the *Kind v. Clif* litigation, if the difference between the p/d and p/d^u confusion rates is not appreciable, then that, standing alone, is *not* sufficient to rule out an inference of actionable likelihood of confusion. The reason, again, is that such a conclusion would inappropriately ignore the possibility that some respondents who are confused by the presence of the unprotectable element p_u might also be confused by the presence of the protectable element p_p. The latter type of confusion *is* actionable and would render confusion due to the joint presence of p_u and p_p actionable.

To deal with this possibility, we add control d^p of the form (p_0, d_1, p_p, d_3); in other words, this control shares inevitable features and protectable features with the plaintiff's mark. Note that this is analogous to d^p in the simplified world of Section II.B.1. As in the simplified version, if the confusion between p and d^p is not appreciable, then that fact, combined with the fact that the difference between the p/d and p/d^u confusion rates is not appreciable, implies that we cannot infer actionable likelihood of confusion. But in the simplified version, we also said that if the confusion between p and d^p *is* appreciable, then we can infer actionable likelihood of confusion because we would know that the protectable element p_p is responsible for appreciable confusion (recall that p_p is a protectable feature like the energy bar's overall configuration). The same is no longer true because not all of the confusion between p and d^p is necessarily attributable to p_p; some of it may be attributable to p_0 alone, which would not be actionable.

Here, the control $d^0 = (p_0, d_1, d_2, d_3)$, which we used in Case 2, comes in handy again. If subtracting the p/d^0 confusion rate from the p/d^p confusion rate leaves an appreciable rate of confusion, then we are justified in inferring actionable likelihood of confusion because we would know that p_p is responsible for appreciable confusion between p and d. However, if the difference between the p/d^p and p/d^0 confusion rates is not appreciable, then we cannot conclude that p_p is responsible for appreciable confusion, which, along with the earlier finding that the difference between the p/d and p/d^u confusion rates is not appreciable, would lead us to conclude that we cannot infer actionable confusion between p and d. The following flowchart summarizes the analysis, and a concrete sample survey employing our methodology is coming in Part III.

Figure 11



As with Section II.B.1, our analysis in this Section has addressed the determination of whether actionable likelihood of confusion can be inferred and not the issue of estimating the rate of actionable confusion. The latter issue is taken up in the Appendix.

C. *Capturing Causation*

In describing our contribution to trademark survey design, we have implicitly relied on a view of causation that, though standard, is not incontestable. The current best practice implicitly relies on an opposing view—but without defending or, for that matter, even articulating it. In this Section, we flesh out and defend our view of causation. Although this causation issue has been deeply thought about in other areas of law, we are the first to discuss it in the trademark context.

Our approach counts three types of people toward the rate of confusion: (1) people confused by p_u and p_p together but not by either alone (the "white balls" in Figure 9); (2) people confused by p_p alone but not by p_u alone (the "dotted balls"); and (3) people confused by either p_u alone or p_p alone (the "striped balls"). But it is not obvious that this third group of people should count. Is it fair to say that the similarity in protectable features *causes* a person's confusion if that person would be confused by the similar *un*protectable features—even if the protectable features were absent? In other words, should

the similarity in protectable features count as a cause if it is a *sufficient* cause but not a *but-for* cause?[167]

We argue that this third group of people should count toward the confusion rate, although we recognize serious arguments in the other direction. But-for causation is the primary way of assessing factual cause in tort law: Conduct is the factual cause of some harm if "the harm would not have occurred absent the conduct."[168] But tort law also generally deems conduct a factual cause of harm if it is one of *multiple sufficient causes*: Each action that alone would have been a but-for cause at the time of the injury is a factual cause under the law.[169] For example, say the defendant negligently starts a fire, which heads toward the plaintiff's property. A fire of unknown origin also makes its way toward the plaintiff's property. The fires merge and, together, destroy the plaintiff's property. If each fire was, on its own, sufficient to destroy the property, the defendant will be held liable for the damage, even though the defendant's negligence was not a but-for cause of the harm.[170]

The same idea should apply here: If the protectable element is sufficient to confuse a person, that person should count toward the rate of confusion, even if the unprotectable element also suffices to confuse that person. For if we required the protectable element to be a but-for cause, then a defendant whose packaging uses *only* the protectable elements *would* be liable, whereas a defendant whose packaging uses *both* the protectable and unprotectable elements would *not* be, assuming the unprotectable elements were also sufficient to cause roughly the same level of confusion. But that would be anomalous. An otherwise-infringing defendant should not be able to immunize itself against liability by adding *more* features in common with the plaintiff's trademark.

Although most authorities would accept multiple sufficient causes as factual causes in situations like the two-fires case, this theory of causation—particularly where one cause is innocent—is not undisputed. The central argument against holding the defendant liable is that the defendant's conduct has made the plaintiff no worse off than it would have been otherwise, so allowing recovery would put the plaintiff in a better position than it would have been absent the

167. For this question to arise, it must be the case not only that either set of features alone would cause confusion but also that the level of confusion caused by both sets of features together is *not appreciably higher* than the level of confusion caused by the unprotectable features alone. If both sets of features together *did* cause appreciably more confusion than the unprotectable features alone, then there is no question that the conduct is actionable: The protectable features are a but-for cause of an appreciable amount of confusion. The current best approach would capture these people in its confusion rate.

168. *See* RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 26 (AM. L. INST. 2010).

169. *See id.* § 27; RESTATEMENT (THIRD) OF TORTS: LIAB. FOR ECON. HARM § 11 cmt. a (AM. L. INST. 2020).

170. *See* Anderson v. Minneapolis, St. P. & S.S.M. Ry. Co., 179 N.W. 45, 46, 49 (Minn. 1920), *overruled on other grounds by* Borsheim v. Great N. Ry. Co., 183 N.W. 519 (Minn. 1921); Michael D. Green, *The Intersection of Factual Causation and Damages*, 55 DEPAUL L. REV. 671, 685–86 (2006).

defendant's conduct.[171] If the purpose of tort law is to return the plaintiff to the position it would have occupied but for the wrong, permitting liability under these circumstances gives the plaintiff too much.

The equivalent objection in the trademark context would be that we should not punish the defendant if its additional taking of the protectable elements does not create appreciably more confusion than would have been created by taking the unprotectable elements alone, which unquestionably would not be actionable. If the purportedly wrongful part of defendant's conduct is not doing appreciably more harm than perfectly lawful conduct, then we should not hold the defendant liable and put the plaintiff in a better position than if the defendant had appropriated only the unprotectable features of the mark.

The argument *for* multiple sufficient causes is, then, "This plaintiff is no *better* off than if the defendant had acted only *unlawfully*." The argument *against* multiple sufficient causes is, "This plaintiff is no *worse* off than if the defendant had acted only *lawfully*." So, the question of which argument is stronger can be framed in terms of which comparison group is more appropriate: the hypothetical unlawfully acting defendant or the hypothetical lawfully acting defendant?

We think that from a policy perspective, it is appropriate to focus on the conduct the law is concerned with: the unlawful conduct, or what might be called "wrongful" conduct. Therefore, the argument *for* multiple sufficient causes is stronger. In tort law, the wrongful conduct is commonly negligence. In the trademark context, it is the appropriation of protectable elements of another's trademark in a way that would create consumer confusion.[172] Just as in the two-fires case the defendant's fire is unquestionably wrongful, in the trademark context the defendant's inclusion of an appreciably confusing protectable element is unquestionably wrongful.

One could argue that including the protectable element is wrongful *only* if the inclusion actually creates—or *causes*—an appreciable likelihood of confusion; if it does not, the defendant has not done anything wrong. Therefore, the analysis in the preceding paragraph begs the question: It assumes the wrongfulness of defendant's conduct, which is the very question to be answered. We think this argument misses the mark. It is true that defendant's taking certain elements of plaintiff's trademark—even taking certain *protectable* elements of plaintiff's trademark—is not, standing alone, a trademark violation. But taking protectable elements *in a way that would by itself create appreciable*

171. Robert J. Peaslee, *Multiple Causation and Damage*, 47 HARV. L. REV. 1127, 1130 (1934); RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 27 reporters' note cmt. d.

172. Of course, to give a full justificatory account, one must go one step further and say *why* consumer confusion is a bad thing (after all, in the absence of trademark law, there would be no consumer confusion resulting from similar trademarks). That deeper reason, as we have discussed, is to reduce consumer search costs and to provide incentives for producers to invest in making better products (at least according to the dominant economic account). *See supra* notes 27–29 and accompanying text. But, regardless of the *ultimate* policy purpose of trademark law, there is universal agreement that the *intermediate* goal is to prevent consumer confusion.

confusion is a violation. To hold otherwise—to require that the conduct actually cause appreciable confusion for the act to be wrongful—would make the definition of the wrong dependent on causation. The same causation analysis, then, would occur twice in the claim, which would be highly anomalous. In tort law, breach of duty and causation are two separate elements, and they should be separate here too.

Of course, wrongful conduct alone is not ordinarily sufficient for liability; the point of the cause-in-fact requirement is to ensure that there is no "negligence in the air"—that wrongfulness ripens into harm before liability is imposed.[173] But when harm *has* occurred, immunizing the defendant on the grounds that the causal link is contestable due to a simultaneous innocent event seems to deviate from the policy purposes of the law. In this context, from a policy standpoint, a windfall to the defendant is a greater anomaly than a windfall to the plaintiff.[174] We have unquestionably wrongful conduct and unquestionable harm—consumers confused about what they are buying, a plaintiff losing business—and we know that the wrongful conduct is one of the contributors to the harm and, indeed, that the wrongful conduct alone would have caused the harm. To the extent the law is primarily concerned with *deterring* undesirable conduct, the case for liability seems as strong here as in the absence of the simultaneous innocent cause. The same is true if the law takes a deontological perspective concerned with recognizing and redressing wrongs: If the defendant has wronged the plaintiff, that entails an obligation to right the wrong.[175]

It is true that constructing this counterfactual—conceiving of the defendant who acts only "unlawfully" or "wrongfully"—is more complicated than we have suggested. The defendant has not taken two separate actions, one wrongful and one permissible. Instead, the defendant has taken both protectable and unprotectable elements from the plaintiff's trademark and has included them together in the same product. The proper analogy, therefore, is not one where the same defendant separately starts one innocent fire and one wrongful fire.[176]

---

173. Martin v. Herzog, 126 N.E. 814, 816 (N.Y. 1920) (quoting FREDERICK POLLOCK, THE LAW OF TORTS 472 (10th ed. 1916)).

174. As Learned Hand put it, "[W]hen one of the two contributing factors is not the result of an actionable fault [but the other factor is], the single tortfeasor cannot be allowed to escape through the meshes of a logical net. He is a wrongdoer; let him unravel the casuistries resulting from his wrong." Navigazione Libera Triestina Societa Anonima v. Newtown Creek Towing Co., 98 F.2d 694, 697 (2d Cir. 1938).

175. Other doctrines in tort law cannot be justified by putting the plaintiff in the position they would have been in without the wrong. For example, the collateral source rule holds that any compensation the injured party receives from a source independent of the tortfeasor (for example, from insurance) is not deducted from the damages the tortfeasor must pay. Rebecca Levenson, Comment, *Allocating the Costs of Harm to Whom They Are Due: Modifying the Collateral Source Rule After Health Care Reform*, 160 U. PA. L. REV. 921, 923 (2012). Here, too, the purposes of the law dictate that a windfall to an injured plaintiff is preferable to a windfall to a wrongdoing defendant.

176. In that analogy, we expect that people's intuitions about liability would depend on the temporal sequence of defendant's actions. Most people, we think, would feel differently about the following two cases: (1) Defendant starts an innocent fire; after doing so, and perhaps

Rather, the proper analogy is one where the same defendant *with a single stroke of the arm* simultaneously starts two fires, one wrongful and one innocent.[177]

However, we believe the difficulty in constructing these counterfactuals cuts in *favor* of liability. In the situation where a single act of the defendant simultaneously starts an innocent and a wrongful fire, absolving the defendant of liability requires not simply imagining a counterfactual where the defendant does not act at all or the defendant takes one action but not another. Instead, crafting a liability-defeating counterfactual involves imagining that the defendant

---

realizing that the fire is destined to burn down plaintiff's home, defendant negligently throws his unextinguished match into a haystack and starts another fire; the fires merge and burn down plaintiff's home; (2) Defendant negligently throws his unextinguished match into a haystack and starts a fire; after doing so, and perhaps realizing that the fire is destined to burn down plaintiff's home, defendant starts an innocent fire; the fires merge and burn down plaintiff's home.

177. Perhaps an even stronger analogy is a discrimination case where the employer takes an adverse employment action for both a legally permissible reason (e.g., the plaintiff's tardiness) and a legally impermissible reason (e.g., the plaintiff's race). These "mixed motive" discrimination cases mirror our trademark example in that: (1) both a permissible influence and an impermissible influence act on someone's mind (the defendant's in the discrimination case, a consumer's in the trademark case); (2) the effect of these influences is simultaneous (the decision to fire, say, in a discrimination case, confusion in the trademark case); (3) both the permissible and impermissible influences originate with a single defendant; and (4) either the permissible influence or the impermissible influence would suffice to cause the harm (termination or confusion). So, we again have multiple sufficient causes, but here the defendant is responsible for *both* the innocent and the wrongful causes, and they manifest in a single action.

Employment law treats these "mixed motive" cases differently depending on which statute provides the cause of action. Title VII of the Civil Rights Act provides that a plaintiff can establish an unlawful employment practice by demonstrating "that race, color, religion, sex, or national origin was a motivating factor for any employment practice, even though other factors also motivated the practice." 42 U.S.C. § 2000e-2(m). However, if the defendant shows that the impermissible consideration was not a but-for cause of the adverse action, the plaintiff's remedy is limited to declaratory relief, injunctive relief (but not reinstatement, hiring, or promotion), and attorney fees and costs. 42 U.S.C. § 2000e-5(g)(2)(B). The Age Discrimination in Employment Act and the retaliation provisions in the Civil Rights Act do not include the same "motivating factor" language. In two closely divided decisions, Gross v. FBL Fin. Servs., Inc., 557 U.S. 167, 176–78 (2009), and Univ. of Tex. Sw. Med. Ctr. v. Nassar, 570 U.S. 338, 360 (2013), the Supreme Court has held that under those statutes, successful plaintiffs must prove but-for causation, and federal courts of appeals have cited those decisions to require but-for causation under, for example, the Americans with Disabilities Act. *See* Natofsky v. City of New York, 921 F.3d 337, 349 (2d Cir. 2019). Justice Kennedy's opinion on the retaliation statute cited the traditional but-for causation standard in tort law to support the view that "because" means "but for." *Nassar*, 570 U.S. at 346–47. Dissenting from the Court's decision on retaliation, Justice Ginsburg cited tort law's embrace of multiple sufficient causes. *Id.* at 383–84 (Ginsburg, J., dissenting).

We agree with the Justices that it is appropriate to draw from tort law to understand causation in statutory contexts such as employment discrimination and trademark law. The analogy holds, even though—unlike in the classic tort examples—the defendant is responsible for both the innocent and the wrongful causes, the causes simultaneously act on a person's mind, and they result in a single action. And we agree with Justice Ginsburg that modern tort law *does* permit recovery in multiple-sufficient-causes cases, so it is sensible to interpret a statutory requirement of causation to encompass sufficient causation. That said, the discrimination analogy does not work entirely in our favor: Five-member majorities have concluded that "because of" means "but for" in the usual course. *Gross*, 557 U.S. at 176–78; *Nassar*, 570 U.S. at 346–47.

acted in a specific way that includes all of the innocent components of its action with none of the wrongful components. There is, in fact, no knowing what product the defendant would have made if it had not included a protectable element of plaintiff's trademark—it might have used different unprotectable elements entirely, or it might have appropriated other protectable elements—but denying liability requires imagining the counterfactual most favorable to the defendant. It is unclear why we should engage in such a level of creative reconstruction to protect a defendant who, after all, has done something wrongful. Even in the typical multiple-sufficient-cause context where the different causes are separate acts done by different persons, the standard answer is that the burden of the entanglement of causes should fall on the wrongdoer: "He is a wrongdoer; let him unravel the casuistries resulting from his wrong."[178] Where, as here, the causes are not clearly separable, the case for liability is even stronger. Therefore, where the defendant's appropriation of the plaintiff's protectable features is one of multiple sufficient causes of confusion, it should be understood to cause confusion.

For the foregoing reasons, we conclude that if the defendant's inclusion of a protectable element similar to plaintiff's is sufficient to confuse a person, then that person should count toward the rate of confusion, even if the defendant has also included a similar unprotectable element that also suffices to confuse that person. At a minimum, we have shown that the current best approach presupposes one answer (the heterodox answer) to a contested question of causation—and it does so without so much as acknowledging that there is a contested question here to be worked out. We hope that even readers who are not persuaded by our answer to the causation question, and hence our multiple-control survey methodology, are at least persuaded that we have shone some light on an important question that existing approaches have ignored.

III. A SAMPLE SURVEY CAPTURING THE MISSING ELEMENT

We have argued that a likelihood of confusion should result in liability only if the confusion arises from similarities between the defendant's trademark and *protectable* elements of the plaintiff's trademark. We have then described a survey methodology that could properly test whether confusion between two trademarks stems from the protected elements of the plaintiff's trademark. We now demonstrate the feasibility of our proposal by presenting an example survey that parties could use as proof that the linkage element is or is not satisfied. Using several sets of hypothetical results, we show how the results should be analyzed to evaluate the likelihood of confusion.

The survey is a modification of the survey conducted in our motivating example case, *Kind v. Clif*,[179] in which Kind claimed that Clif's Mojo Bar packaging infringed its Kind Bar trade dress. As noted above, the plaintiff's expert used a control that did not contain unprotectable elements common

178. *Navigazione Libera*, 98 F.2d at 697.

179. Kind LLC v. Clif Bar & Co., No. 14 Civ. 770, 2014 WL 2619817 (S.D.N.Y. June 12, 2014).

to the plaintiff's trade dress and the defendant's wrapper (Figure 5).[180] The defendant's expert then modified the survey, using controls that *did* contain those unprotected elements (Figure 6).[181] But neither survey used a control that could determine whether the common *protected* elements were sufficient to cause the survey respondents' confusion.

Our survey adds controls that allow us to account for this possibility and home in on the relevant question: Does the defendant's product create a likelihood of confusion because it uses protected elements of the plaintiff's trade dress? In the terminology of Part II (restated in the Notation Table in Section C of the Appendix), if the plaintiff's trade dress takes the form $p = (p_0, p_u, p_p, p_3)$ and the defendant's trade dress takes the form $d = (p_0, p_u, p_p, d_3)$,[182] we include not only the control $d^u = (p_0, p_u, d_2, d_3)$, which the defendant's expert included, but also the controls $d^p = (p_0, d_1, p_p, d_3)$, which neither expert used, and $d^0 = (p_0, d_1, d_2, d_3)$, which the plaintiff's expert used.

Trademark surveys typically come in two varieties: the "Eveready" format and the "Squirt" format.[183] An Eveready[184] survey does not show respondents the plaintiff's product. Instead, respondents see only the defendant's product, and the survey asks an open-ended "source confusion" question, such as "Who makes or puts out this product?"[185] That question is often followed by "sponsorship confusion" or "affiliation confusion" questions that ask whether the respondent believes the product manufacturer is sponsored or approved by or affiliated with another company and, if so, which company.[186]

The Eveready format is widely accepted by courts and commentators.[187] But if the mark is not "top-of-mind," such that it will come to mind whenever a consumer encounters a similar mark, an Eveready survey may underestimate

180. *See* Mantis Declaration, *supra* note 119, ¶ 3.

181. *See* Rappeport Declaration, *supra* note 123, at 6–7.

182. Recall that p_0 represents "inevitable" features such as the type of product (energy bar) and the fact that it comes in packaging; p_u represents unprotectable features of the plaintiff's trade dress that are in common or similar with the defendant's trade dress, such as see-through wrapping; p_p represents protectable features of the plaintiff's trade dress that are in common or similar with the defendant's trade dress, such as the overall arrangement of various features of the energy bar package; and p_3 represents other features of the plaintiff's trade dress that are not shared with the defendant's trade dress, such as the Kind logo.

183. 5 MCCARTHY, *supra* note 5, § 32:173; Swann, *supra* note 110, at 53.

184. The name "Eveready" comes from the mark in the Seventh Circuit case in which the format was endorsed. Union Carbide Corp. v. Ever-Ready Inc., 531 F.2d 366, 381 (7th Cir. 1976); *see* Swann, *supra* note 110, at 53.

185. Swann, *supra* note 110, at 53; 5 MCCARTHY, *supra* note 5, § 32:174.

186. Swann, *supra* note 110, at 56–58.

187. *Id.* at 58, 62 (calling the Eveready format "the gold standard" for "cases involving top-of-mind marks"); 5 MCCARTHY, *supra* note 5, § 32:174.50.

the likelihood of consumer confusion.[188] When the two products are placed near each other in the marketplace (e.g., on the shelf in a store), a consumer seeing *both* products may believe they are affiliated, even if the defendant's product alone does not cause a consumer to think of the plaintiff's product.[189] For non-top-of-mind products that are in proximity in the marketplace, then, a Squirt survey may be more appropriate.

A Squirt[190] survey presents the respondent with both the plaintiff's mark and the defendant's mark.[191] In a traditional Squirt survey, respondents are shown the two marks and asked whether they believe the two products are put out by the same company and why.[192] But this format can lead to substantial "demand effects."[193] Demand effects—a long-recognized concern in social science experiments[194]—occur when respondents infer the purpose of the survey and answer questions in a way that satisfies it.[195] Large demand effects can severely bias survey results.[196] While good controls can help researchers account for these effects, they can still be so large as to render the survey "inherently suspect."[197] So more recent Squirt surveys use various methods to make the survey less leading, including presenting subjects with a product "lineup" and asking which, if any, of the products were made by the same company.[198] Squirt surveys have sometimes been criticized by commentators and rejected by courts for being improperly leading.[199]

---

188. Jerre B. Swann, *Eveready and Squirt—Cognitively Updated*, 106 TRADEMARK REP. 727, 733–35 (2016). For example, one of us purchases goat cheese on a weekly basis, always taking it from the same shelf at Fairway Market. She is pretty sure she buys the same brand every week—the label is mostly white and blue—but she does not keep the product's trade dress top-of-mind. If she saw a different type of cheese with a similar label, she might not spontaneously identify it with her goat cheese. However, if the two cheeses were next to each other at Fairway, she might well assume they were affiliated. *See id.* at 733 (identifying "routinely purchased grocery items" as goods that "may not enjoy high levels of unaided recall because" the consumer can recognize them based on location).

189. *See id.* at 734–37.

190. The Squirt format is, similarly, named after the mark in its case of origin, *SquirtCo v. Seven-Up Co.*, 628 F.2d 1086, 1089 n.4, 1090–91 (8th Cir. 1980). *See* Swann, *supra* note 110, at 53.

191. 5 MCCARTHY, *supra* note 5, § 32:174.50.

192. Swann, *supra* note 110, at 64. "Sponsorship confusion" and "affiliation confusion" questions may follow. *See id.*

193. *Id.* at 65.

194. *See* Martin T. Orne, *On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications*, 17 AM. PSYCH. 776, 779 (1962).

195. Itamar Simonson & Ran Kivetz, *Demand Effects in Likelihood of Confusion Surveys: The Importance of Marketplace Conditions*, *in* TRADEMARK AND DECEPTIVE ADVERTISING SURVEYS: LAW, SCIENCE, AND DESIGN, *supra* note 110, at 243, 243.

196. *Id.* at 244.

197. Swann, *supra* note 110, at 65.

198. 5 MCCARTHY, *supra* note 5, § 32:177; Rappeport, *supra* note 139, at 236–37 (describing "array" surveys).

199. 5 MCCARTHY, *supra* note 5, § 32:174.50.

Because Kind may not be a sufficiently top-of-mind mark for the Eveready method, and because snack bars are often presented next to each other in the market,[200] we use a variant of the Squirt survey, like the survey in *Kind v. Clif*.[201] Respondents are shown three different products in sequence: first, five packs of Trident chewing gum, then five Kind Bars, then five packs of Tic Tac mints. (We include the gum and mints only to remove focus from the snack bars.) They are then shown three additional products in sequence and asked, for each one, if it was made with the approval of the company that made the earlier product of the same type. Half of the subjects see the original Trident gum and Altoids mints, while half see Orbit gum and the original Tic Tac mints.[202]

More importantly, one quarter of the respondents see the defendant's allegedly infringing Mojo Bar (d), one quarter see a bar that contains prominent but unprotectable features of the plaintiff's packaging (d^u), one quarter see a bar that contains protectable features of the plaintiff's packaging, without the prominent but unprotectable features (d^p), and one quarter see a bar that contains only the "inevitable" features of the plaintiff's packaging (d^0). Although the plaintiff's expert in *Kind v. Clif* included d^0 and the defendant's expert included d^u as a control, only by using *both* controls *and* d^p, which neither party's expert used, can we determine whether any confusion can be attributed to the defendant's use of the protectable features of the plaintiff's trade dress.

Although this design runs the risk of demand effects, we think it is appropriate for several reasons. First, our controls will allow us to account for demand characteristics in our analysis by subtracting out the demand-affected confusion rate for those controls. Second, the use of products other than the bars at issue should "remov[e] focus away from the snack bar products tested."[203] Third, as the plaintiff's expert notes, a similar method has been accepted by courts in the past.[204] Fourth, the basic methodology was accepted by both the plaintiff's expert and the defendant's expert in the *Kind v. Clif* litigation.[205]

200. See, for example, the second floor of the Fairway on Broadway and 74th Street.

201. *See* Mantis Declaration, *supra* note 119, ¶¶ 12–14.

202. The plaintiff's survey in *Kind v. Clif* showed the respondents one of the same non-snack-bar products that they saw in the first half to remove focus from the snack bars. *Id.* ¶ 13. We are concerned that this could bias results in favor of the defendant by tempting some participants to spot the products that differ and label them unrelated. Still, because our intervention concerns the controls, we use this feature of the plaintiff's survey.

203. *Id.*

204. *See, e.g.*, Kraft Gen. Foods, Inc. v. Allied Old Eng., Inc., 831 F. Supp. 123, 130–31 (S.D.N.Y. 1993) (crediting a study in which products of various types were presented on a shelf, including either the defendant's product or a control, and then respondents were asked in sequence whether several products, including plaintiff's, were made by the same company that made the product of equivalent type they had seen on the shelf).

205. *See* Rappeport Declaration, *supra* note 123, at 4.

A. THE SURVEY INSTRUMENT

An annotated survey instrument follows. Questions in boxes will be presented to respondents; annotations are outside the boxes.

The survey will begin with informed consent:

> You are invited to participate in a research survey about consumer perceptions.
>
> The survey should take less than 10 minutes to complete, and your responses will be anonymous. Participation in this survey is voluntary and has no foreseeable risks. You may stop at any time. You must be 18 years or older and live in the U.S. to participate in this survey. Please use a desktop, laptop, or large tablet to take the survey. Do not use a phone.
>
> If you have any questions about this survey, please contact [survey expert name and contact information].

The survey collects basic demographic information:

> **Q1.** In which state do you live?
> - [Dropdown of all states & D.C. + not in U.S.]
>
> **Q2.** What is your gender?
> - Male
> - Female
> - Non-binary
>
> **Q3.** What is your age?
> - Under 18
> - 18 to 24
> - 25 to 34
> - 35 to 44
> - 45 to 54
> - 55 to 64
> - 65 or older

Respondents who do not live in the U.S. or are under 18 will be dropped.

We confine our sample to an appropriate universe by determining whether the participants are consumers of snack bars.

> **Q4.** In the past 3 months, have you purchased any of the following products?
> - Chewing gum (Y/N)
> - Mints (Y/N)
> - Snack bars (Y/N)
>
> **Q5.** In the next 3 months, are you likely to purchase any of the following products?
> - Chewing gum (Y/N)
> - Mints (Y/N)
> - Snack bars (Y/N)

Respondents who answer "no" to both snack bar questions will be excluded from the sample. A pilot survey would help determine the percentage of respondents likely to be excluded on these grounds.

We next eliminate respondents ineligible because they are in an industry that might be too familiar with the products or branding.

> **Q6.** Do you or any members of your household currently work for any one of the following types of companies? (Choose one answer per row)
> - An advertising agency, sales promotion firm, public relations firm, or marketing research organization (Y/N/DK)
> - A manufacturer or distributor of any food products (Y/N/DK)
> - A retail store or any other establishment where you personally sell food products (Y/N/DK)

We exclude respondents who answer "yes" or "don't know" to any of these questions.

We then provide instructions for taking the survey, taken from the instructions in the plaintiff's survey in *Kind v. Clif*, followed by an attention check of our own.

Before continuing with this survey, please carefully read the following instructions:

1. Please take the survey in one session.
2. While taking this survey, please do not at any time open any other windows or tabs on this computer or any other computer.
3. While taking this survey, please do not at any time use any hand-held electronic device, including a phone.
4. Please do not view any other written material or electronic devices while taking this survey.
5. Please do not consult or talk with any person while taking this survey.
6. Please do not disclose to any party or copy or download any information that you will see in this survey.
7. Please wear any corrective lenses that you normally wear while using a computer.

Q7. Do you understand the above instructions and agree to follow them?

- Yes
- No
- Don't know/Not sure

Q8. Please show you are paying attention by selecting "knees" below:

- Head
- Shoulders
- Knees
- Toes

Respondents who do not respond "Yes" or who fail the attention check will be excluded. Those who have made it past this point will continue to the substantive portion of the survey.

> Now you will see three photos, one at a time. The first photo shows several packages of the same brand of chewing gum. The second photo shows several packages of the same brand of snack bars. And the third photo shows several packages of the same brand of mints.
>
> Please look at the products shown in each photograph as you would if you were thinking about buying them.
>
> When you are done looking at each photo, please click the "Continue" button to advance to the next photo.

The respondent will then see three photos, as described above. The "snack bars" will be plaintiff's Kind Bars.

Chewing Gum:

Snack Bars:

Mints:

The respondent then proceeds to the portion of the survey that measures confusion.

> Now you will see three more photos, and then you will be asked some questions. For each question, if you don't know or don't have an answer, please don't guess. Just indicate you don't know or don't have an answer and proceed to the next question.

The order in which the products are presented will be randomized. The respondents will be split into eight (two times four) groups. Half of the respondents will see the same chewing gum as before (not shown below) and different mints, while half will see a different chewing gum and the same mints (not shown below).

Our intervention comes in the four-group division:

- One quarter of subjects will see d, the defendant's allegedly infringing product: the Clif Mojo Bar.
- One quarter of subjects will see d^u, a bar that contains prominent but unprotectable features of the plaintiff's packaging. Like defendant's expert in the actual litigation, we use the TrueBar, which has transparent wrapping—a functional, unprotectable feature of the Kind Bar—but otherwise has a different design.
- One quarter of subjects will see d^p, a bar that contains only protectable features of the plaintiff's packaging (and, inevitably, the "inevitable" features). Here, we modify defendant's Mojo Bars, replacing the transparent portion with an enlarged picture of one or more ingredients in the snack bar.
- Finally, one quarter see d^0, a bar that contains only the inevitable features of the plaintiff's packaging. Here, that is an earlier design of the Mojo Bar packaging that does not contain design elements of the Kind Bar wrapper.

For each picture the respondent sees, they are asked three questions. First, "Do you think this brand of [product type] IS or IS NOT made by, or made with the approval or sponsorship of, the same company that makes the [product type] you saw in the earlier photo?" We follow the plaintiff's expert in *Kind v. Clif* in not separating source confusion from sponsorship confusion. Either one would indicate the necessary confusion. But a study that did separate the two could use the same basic methodology. Second, the subject is asked, "What makes you say that?" Finally, the subject will be asked "Anything else?" to allow them to provide any additional information.

Q9. Shown in this photo are five packages of chewing gum:

Do you think this brand of chewing gum IS or IS NOT made by, or made with the approval or sponsorship of, the same company that makes the chewing gum you saw in the earlier photo?

- Yes/IS
- No/IS NOT
- Don't know/No opinion

Q10. What makes you say that?

- [Text box]

Q11. Anything else?

- [Text box]

Q12. Shown in this photo are five packages of mints:

Do you think this brand of mints IS or IS NOT made by, or made with the approval or sponsorship of, the same company that makes the mints you saw in the earlier photo?

- Yes/IS
- No/IS NOT
- Don't know/No opinion

Q13. What makes you say that?

- [Text box]

Q14. Anything else?

- [Text box]

Condition *d*, the defendant's product:

Q15. Shown in this photo are five packages of snack bars:

Do you think this brand of snack bars IS or IS NOT made by, or made with the approval or sponsorship of, the same company that makes the snack bars you saw in the earlier photo?

- Yes/IS
- No/IS NOT
- Don't know/No opinion

Q16. What makes you say that?

- [Text box]

Q17. Anything else?

- [Text box]

Condition d^u, pictures of a snack bar with a transparent section of its wrapper, a prominent but unprotectable feature of the plaintiff's trade dress. The remainder of the TrueBar's trade dress differs from the Kind Bar, so it is unlikely to cause confusion on protected elements. This is akin to the control that the defendant's expert used in *Kind v. Clif*:

Q15. Shown in this photo are five packages of snack bars:

Do you think this brand of snack bars IS or IS NOT made by, or made with the approval or sponsorship of, the same company that makes the snack bars you saw in the earlier photo?

- Yes/IS
- No/IS NOT
- Don't know/No opinion

Q16. What makes you say that?

- [Text box]

Q17. Anything else?

- [Text box]

Condition d^p, pictures of the defendant's product, where the feature overlapping with the *unprotected* part of the plaintiff's trade dress has been removed, so only the protected elements remain. Here, the transparent portion of the defendant's wrapper has been modified to instead display enlarged pictures of ingredients.[206] Neither the plaintiff's expert nor the defendant's expert used a similar control in *Kind v. Clif*:

Q15. Shown in this photo are five packages of snack bars:

Do you think this brand of snack bars IS or IS NOT made by, or made with the approval or sponsorship of, the same company that makes the snack bars you saw in the earlier photo?

- Yes/IS
- No/IS NOT
- Don't know/No opinion

Q16. What makes you say that?

- [Text box]

Q17. Anything else?

- [Text box]

206. The image was modified using GIMP, an open-source raster graphics editor. *See* GIMP, https://www.gimp.org [https://perma.cc/94NQ-MKVW].

Condition d^0, a picture of snack bars that share only non-trade-dress features with the plaintiff's product. Here, we use an earlier version of the defendant's Mojo Bars, which are rectangular snack bars in wrappers with flavors printed on the front, but they otherwise share few features with the plaintiff's product. The plaintiff's expert used this control in *Kind v. Clif*:

Q15. Shown in this photo are five packages of snack bars:

Do you think this brand of snack bars IS or IS NOT made by, or made with the approval or sponsorship of, the same company that makes the snack bars you saw in the earlier photo?

- Yes/IS
- No/IS NOT
- Don't know/No opinion

Q16. What makes you say that?

- [Text box]

Q17. Anything else?

- [Text box]

Thank you for completing this survey.

B. Sample Results and Discussion

To illustrate how our proposal would work in practice, this Section offers four sets of hypothetical survey results and discusses how they should be

interpreted.[207] We focus on responses to Question 15,[208] which asks the respondent whether they think a pictured snack bar is or is not made by, or made with the approval or sponsorship of, the same company that makes the snack bar the respondent saw in the earlier photo (the Kind Bar). The snack bar the respondent sees while answering Question 15 might be either: d, the accused Clif Mojo Bar; d^u, the TrueBar, which has a transparent section but different overall configuration; d^p, the modified Clif Mojo Bar with a similar overall configuration of the package but with pictures of fruit in place of a transparent section; or d^0, a previous iteration of the Clif Mojo Bar, which looks nothing like the Kind Bar, except that they are both energy bars in wrappers.

For purposes of this analysis, we assume the court will deem survey-based actionable confusion levels of at least 20% to suggest that an appreciable number of consumers are likely to be confused. In reality, this threshold number will depend on the specifics of the case, including the strength of other evidence of likely confusion, and higher or lower levels of survey confusion may be weaker or stronger evidence.[209] But for our purposes, we assume the court has set a threshold level of 20%: If and only if the survey evidence demonstrates actionable confusion at or above that level will the court deem it supportive of the plaintiff's case.

In each of the following scenarios, we provide some hypothetical survey results and analyze whether the actionable confusion rate reaches this 20% threshold. This analysis maps onto the decision tree in Figure 11. We show how our method of defining actionable confusion compares to both the naïve method, which may overestimate actionable confusion by including confusion arising from unprotectable features, and to current best practices, which may underestimate actionable confusion by excluding people who would be confused by either the protectable features alone or the unprotectable features alone.

207. As we previously noted, current practice in the interpretation of trademark survey results does not show sufficient attention to questions of statistical inference. Future work should address this important issue, but it is beyond the scope of our intervention, so we will set it aside for purposes of interpreting our sample survey results. *See supra* note 141.

208. We do not include survey results for Questions 1 through 14. Questions 4 through 8 determine eligibility. Questions 4 and 5 limit the survey to people who have purchased or are likely to purchase snack bars; Question 6 eliminates people too familiar with relevant industries; Question 7 eliminates anyone who says they do not understand instructions; and Question 8 eliminates subjects not paying sufficient attention to follow a simple instruction. We assume our results below reflect only respondents deemed survey eligible. Other questions provide demographic information, to ensure a representative sample, or serve to prevent demand effects. Questions 16 and 17 ask subjects why they answered how they did. These answers would provide ammunition for survey experts to critique each other's methodology or bolster their conclusions, but they do not figure into our analysis here.

209. *See* 5 MCCARTHY, *supra* note 5, § 32:185; *id.* § 32:188.

1. Clearly Insufficient Confusion

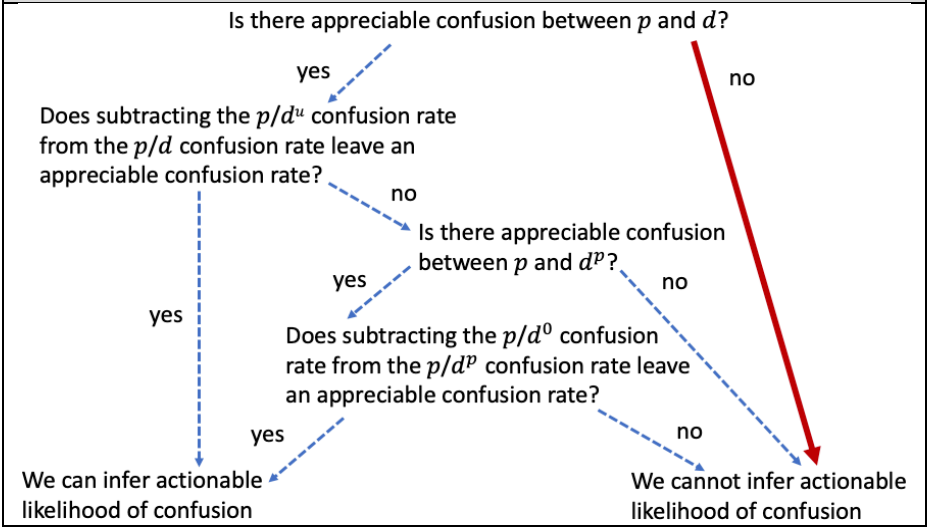
Hypothetical 1:

- In condition d, 15% of respondents answer "Yes/IS" to Question 15.

If fewer than 20% of respondents in condition d say that they think the defendant's product is made by, or made with the approval or sponsorship of, the same company that makes the plaintiff's product, then there is no appreciable confusion *in the aggregate*. If this *raw* confusion rate is below the legal threshold, even before adjusting for controls, then necessarily the *actionable* confusion rate is also below the threshold. It does not matter how many respondents are confused by the controls; there can be no liability. We follow the rightmost branch in the decision tree and conclude that the survey provides no evidence of confusion:

Figure 11a

Decision Tree on Whether to Infer Actionable Likelihood of Confusion from Surveys Where Confusion Due to Inevitable Product Features Is Not Negligible, with the Hypothetical 1 Decision Path Highlighted



2. Our Method and Current Best Practices Both Find Sufficient Confusion

Hypothetical 2:

- In condition d, 35% of respondents answer "Yes/IS" to Question 15.
- In condition d^u, 7% of respondents answer "Yes/IS" to Question 15.

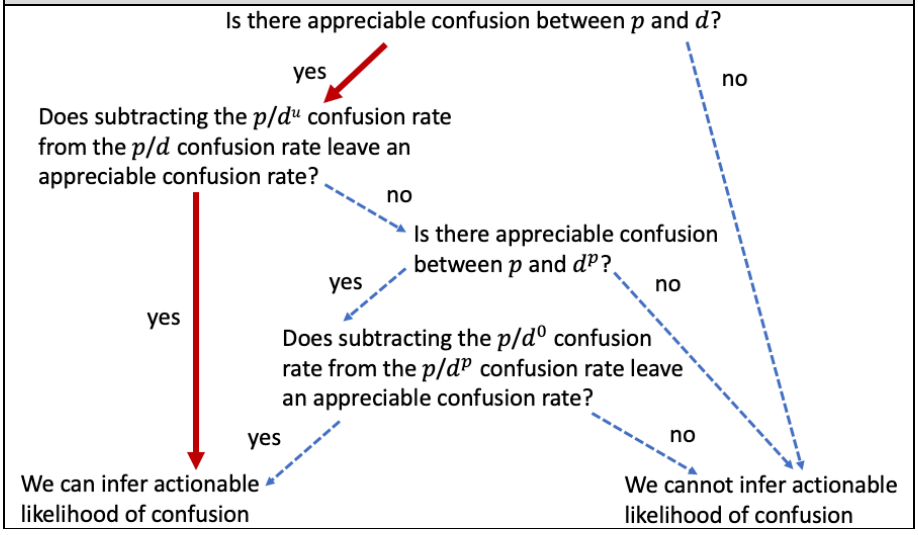
Under both our method and the current best practices, this survey provides evidence of likelihood of confusion. The raw rate of confusion, 35%, far exceeds the threshold of 20%. And only 7% of people are confused by a control that contains only the unprotectable elements of the plaintiff's trade dress, which

includes the "inevitable" features, like the fact that it is approximately the shape and size of a snack bar.[210] This is the type of control used by the best survey experts today. Those experts would conclude that there is sufficient confusion here because subtracting the control rate of confusion from the raw rate of confusion leaves 28%, which exceeds the 20% threshold.

Similarly, under our analysis, we are on the leftmost path of Figure 11:

Figure 11b

Decision Tree on Whether to Infer Actionable Likelihood of Confusion from Surveys Where Confusion Due to Inevitable Product Features Is Not Negligible, with the Hypothetical 2 Decision Path Highlighted



We can infer sufficient actionable confusion here because we know that the sum of (1) the people who are confused by the protectable features alone but not the unprotectable features alone, and (2) the people who are confused by the protectable and unprotectable features in combination but not by either alone, is 28%, greater than the required legal threshold of 20%. Even without accounting for people who would be confused by either the protectable elements alone or the unprotectable elements alone, we have a sufficient rate of confusion.

210. For this reason, it does not matter how many people are confused by d^0, the control containing *only* inevitable features. Under our monotonicity assumption, see *supra* note 125, anyone confused by d^u would also be confused by d^0. So, when we subtract the rate of confusion in the d^u condition, that includes everyone who would be confused by d^0. Were we to *additionally* subtract the rate of confusion in the d^0 condition, that would double-count people and could result in an understated rate of confusion.

3. Our Method Finds Sufficient Confusion, Current Best Practices Would Not

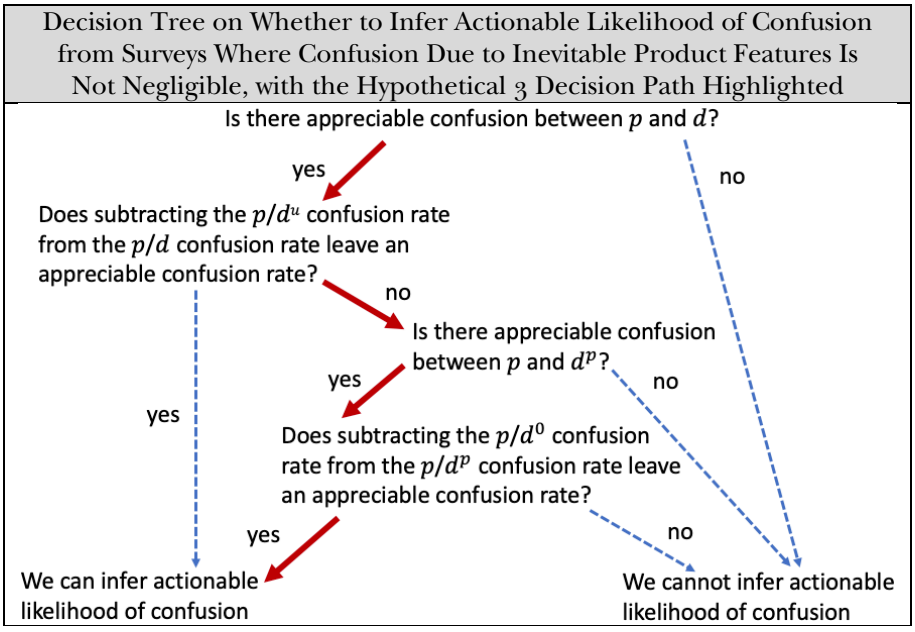
Hypothetical 3:

- In condition d, 30% of respondents answer "Yes/IS" to Question 15.
- In condition d^u, 12% of respondents answer "Yes/IS" to Question 15.
- In condition d^p, 25% of respondents answer "Yes/IS" to Question 15.
- In condition d^0, 2% of respondents answer "Yes/IS" to Question 15.

Here, we see how our method can come to a different conclusion than current best practices. A survey expert using current best practices would perform the same exercise as in Hypothetical 2: They would subtract the rate of confusion in the unprotectable-features control condition (12%) from the rate of confusion in the defendant's-product condition (30%). This rate, 18%, fails to meet the threshold of 20%. Therefore, this expert would conclude, the survey shows insufficient confusion to support the plaintiff.

However, this method neglects any subjects who would be confused *both* by d^u, the unprotectable-features control, *and* by d^p, the protectable-features control. As we have argued, these people should count toward the rate of confusion. Therefore, under our proposal, the analysis would not stop there: We would ask, if we take the rate of confusion in the protectable-features control condition (25%) and subtract the rate of confusion in the inevitable-features control condition (2%), does that number exceed the threshold confusion level? If so, we can be confident that the percentage of people who are confused by the protectable features alone is above the threshold. And if the percentage of people confused by the protectable features *alone* is higher than the threshold, then the actionable confusion rate—which includes both people confused by the protectable features alone *and* people confused by both features together but not by either alone—is necessarily also above the threshold. Here, that condition is satisfied: 25% - 2% = 23%, which exceeds the 20% threshold. We are on the following path in Figure 11:

Figure 11c



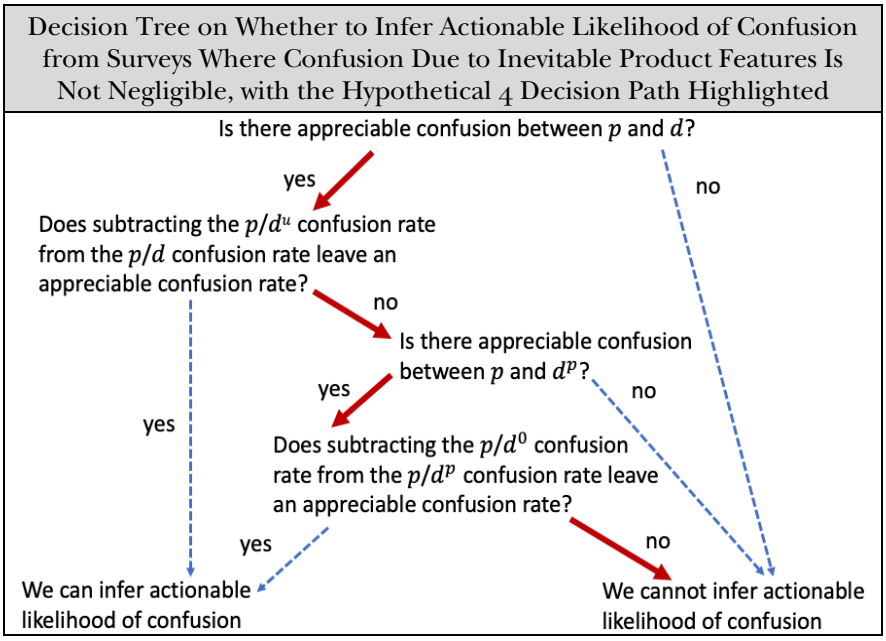
4. Our Method Achieves a Better Estimate than Current Best Practices

Hypothetical 4:

- In condition d, 25% of respondents answer "Yes/IS" to Question 15.
- In condition d^u, 16% of respondents answer "Yes/IS" to Question 15.
- In condition d^p, 21% of respondents answer "Yes/IS" to Question 15.
- In condition d^0, 2% of respondents answer "Yes/IS" to Question 15.

Both current best practices and our method would conclude that the survey results fail to show sufficient confusion. Using the same method as in Hypotheticals 2 and 3, a current survey expert would subtract the rate of confusion in the unprotectable-features control condition (16%) from the rate of confusion in the defendant's-product condition (25%) and get 9%, failing to meet the threshold of 20%. Using the same method as in Hypothetical 3, we would subtract the rate of confusion in the inevitable-features control condition (2%) from the rate of confusion in the protectable-features control condition (21%) and get 19%, also failing to meet the threshold of 20%. Our analysis would follow this path in Figure 11:

Figure 11d



But our method has an advantage over current best practices, even when it comes to the same conclusion, because our method provides a better estimate of the true actionable rate of confusion. Here's why[211]: The true actionable rate of confusion is the sum of (1) people confused by the protectable and unprotectable features together but not by either set of features alone; (2) people confused by the protectable features alone but not by the unprotectable features alone; and (3) people confused *both* by the protectable features alone and by the unprotectable features alone. Under current survey methodology, an expert takes the total rate of confusion (condition d) and subtracts the people confused by the unprotectable features (condition d^u), which equals the sum of groups (1) and (2). Although, as demonstrated in the Appendix, our second control (d^p) does not allow us to calculate the precise sum of all three groups,[212] it allows us to calculate the percent of people confused by the protectable features alone, which is the sum of groups (2) and (3). By taking the *higher* of (1) + (2) and (2) + (3) instead of just (1) + (2), our approach pinpoints a tighter, more precise lower bound for the true actionable rate of confusion.[213]

This improved estimate can have practical consequences. In Hypothetical 4, the confusion rate using existing methods (9%) is so far below the threshold

211. See the Appendix for a more precise explanation.

212. *See infra* note 221 and accompanying text.

213. See the Appendix for a more precise explanation.

that a court might consider it to be evidence *against* likelihood of confusion.[214] But our method shows that the true actionable rate of confusion is *at least* 19%. Although we cannot reliably conclude, based on these responses, that the actionable level of confusion exceeds the 20% threshold, the survey also does not tend to prove that the level of confusion *fails* to meet the threshold. This survey should not count against the plaintiff, who should have the opportunity to prove its case using other methods. Current best methods would erroneously conclude that the plaintiff had a particularly weak case.

In sum, by constructing a survey that incorporates our novel control and analyzing several sets of hypothetical results, we have demonstrated both the feasibility and the practical significance of our intervention.

CONCLUSION

We have argued that the action for trademark infringement, as currently litigated and understood, is missing an element. Trademark validity and likelihood of confusion are not enough; there should also be a requirement to trace the confusion to a protectable feature of the trademark. Otherwise, there can be liability without infringement of a cognizable right. In fleshing out this argument, we have contributed to trademark law on both theoretical–doctrinal and methodological levels. On the theoretical–doctrinal level, we have shown why the proposed causal tracing requirement makes sense by reference to the purposes of trademark law and by analogy to other areas of law, including other intellectual property regimes. On the methodological level, we have shown how to prove the missing element by systematically laying out how to design surveys that get at the cause of confusion. We have explained why the current best practice of incorporating a control survey that comes as close as possible to the accused product without including the allegedly infringing feature is incomplete, and why proper causal tracing requires two or more control conditions, not just one. Our new approach to survey design not only promises more accurate estimates of the actionable rate of confusion; it also grapples with broader problems about the nature of factual causation that were previously ignored in trademark law. Finally, by introducing a realistic survey and interpreting four different sets of hypothetical results, we have shown that our proposed reworking of the elements of trademark infringement is not only theoretically fundamental but also practically doable.

It is worth noting that our doctrinal and methodological contributions favor different parties. At the doctrinal level, we have proposed a requirement for trademark infringement that would help some defendants avoid liability. As we have shown, courts and trademark parties frequently miss this requirement, potentially allowing plaintiffs to win too easily. When they do not miss it—typically in the context of designing and judging trademark surveys in high-stakes cases—the best existing methodology actually overshoots the mark: It ends up overprotecting defendants against liability. Our methodological

214. *See* 5 MCCARTHY, *supra* note 5, § 32:189.

intervention corrects this defendant-friendly error. Our aim with this Article is not to narrow or to expand trademark liability; we only hope our analysis will help courts get it right.

APPENDIX

In this Appendix, we derive an estimate of the actionable rate of confusion. Recall our thesis that not all confusion arising from similarities between the allegedly infringed and allegedly infringing trademarks is actionable. The task is to estimate what part of that confusion is attributable to the protectable elements of the allegedly infringed trademark—that is, to go from the raw confusion rate to the actionable confusion rate. We do this first in the "simplified" context of Section II.B.1, where confusion due to inevitable non-trademark features is assumed to be negligible, and then in the "real world" context of Section II.B.2, where that assumption is relaxed. Though the latter version is more realistic, and thus chosen for our hypothetical survey in Part III, studying the simplified version is sufficient for readers to understand the core insights.

A. INTO THE WEEDS: THE SIMPLIFIED WORLD

Recall the following notation: p, the allegedly infringed trade dress, takes the form (p_u, p_p, p_3) and d, the allegedly infringing trade dress, is of the form (p_u, p_p, d_3). So the defendant's trade dress shares an unprotectable element (p_u) and a protectable element (p_p) with the plaintiff's trade dress, and they each have other elements (p_3 and d_3, respectively) that are not similar. The controls are $d^u = (p_u, d_2, d_3)$ and $d^p = (d_1, p_p, d_3)$.

Denote by R_d the confusion rate between p and d, which we would get from the treatment survey. Likewise, denote by R_{d^u} and R_{d^p} the confusion rates between p and d^u and between p and d^p, respectively, which we would get from the control surveys.

The key to deriving the correct estimate is to partition R_d into disjoint subsets. The following partition works (where, for the benefit of readers who are unfamiliar with set theory, we slightly abuse notation and use uppercase R, with given subscripts, to refer both to sets of respondents and to rates of confusion corresponding to those sets)[215]:

$$R_d = R_1 + R_2 + R_3 + R_4$$

where

- R_1 represents those who are confused both by p_u alone and by p_p alone[216]
- R_2 represents those who are confused by p_u alone but not by p_p alone[217]

215. *See infra* note 217.

216. When we write "confused by p_i," we mean "confused by the similarity between the two trade dresses with respect to the element p_i."

217. To make this Appendix easy to follow for readers who are not familiar with set theory, in the main text we slightly abuse notation and use R_i to refer both to sets of respondents and to

- R_3 represents those who are confused by p_p alone but not by p_u alone
- R_4 represents those who are confused by (p_u, p_p) together but not by p_u or p_p alone

Note that this is indeed a partition—the subsets "add up" exactly to R_d[218] and the subsets are pairwise disjoint, meaning there is no overlap between any two of them.[219]

Our ultimate quantity of interest—the actionable rate of confusion—is not the aggregate (or "raw") confusion rate represented by R_d but rather that portion of R_d which is attributable to the protectable elements of p. Let us call this actionable confusion rate x. With the aid of our partition, we can see that,

$$x = R_1 + R_3 + R_4$$

Now let us explore how the controls can help us estimate the actionable confusion rate x. Given that the only thing p and d^u have in common is p_u, the confusion rate yielded by the first control survey represents all of those who are confused by p_u alone (regardless of whether or not they are also confused by p_p alone). That is,

$$R_{d^u} = R_1 + R_2$$

We are now in a position to see exactly why the current gold-standard approach of subtracting R_{d^u} from R_d to arrive at x is wrong: The subtraction yields $R_3 + R_4$, not $R_1 + R_3 + R_4$, which is what we want. That is a formal statement of our argument in Section II.B.1 that this approach lowballs the actionable rate of confusion, and we can see exactly by how much it falls short (namely, by R_1).[220]

---

rates of confusion corresponding to those sets. A more precise notation would say that r_1 represents the set of respondents who are confused both by p_u alone and by p_p alone, and R_1 represents the cardinality of that set divided by the number of survey respondents. This way, R_1 represents the confusion rate corresponding to r_1, and so on for R_2 through R_4.

218. Formally, $r_1 \cup r_2 \cup r_3 \cup r_4 = r_d$.

219. That is, the intersection between any two of them is empty. Formally, $r_i \cap r_j = \emptyset \ \forall i \neq j$, $i, j \in \{1,2,3,4\}$.

220. It appears that the thought process underlying the current approach does not use the partition we have in mind but rather an alternative partition. The reason it goes wrong is that the "partition" it uses is actually not a partition.

Here is a formal statement of the best survey experts' reasoning (corresponding to the intuitive explanation in Figures 7–9 and accompanying text): $R_d = R_a + R_b + R_c$ where R_a represents those who are confused by p_u alone, R_b represents those who are confused by p_p alone, and R_c represents those who are confused by (p_u, p_p) together but not by p_u or p_p alone. Or, more precisely, r_a, r_b, r_c, r_d represent the sets and R_a, R_b, R_c, R_d represent the confusion rates corresponding to those sets. The quantity of interest, the actionable rate of confusion, is therefore $R_b + R_c$. Now R_{d^u} yields the confusion rate due to p_u alone, which is to say R_a. So subtracting R_{d^u} from R_d yields the actionable confusion rate.

The reason this seemingly correct reasoning is actually incorrect is that its original partition of r_d is not in fact a partition. The subsets r_a and r_b are *not* disjoint. There can be people who are confused *both* by p_u alone *and* by p_p alone. (Formally, $r_a \cap r_b \neq \emptyset$, which implies $R_d \neq R_a + R_b + R_c$.) So R_a can include some people who we should *not* subtract from R_d to arrive at

Now we bring in our second control, d^p. Because the only element shared by p and d^p is p_p , the confusion rate from the second survey represents all of those who are confused by p_p alone (regardless of whether or not they are also confused by p_u alone). That is,

$$R_{d^p} = R_1 + R_3$$

The ultimate question is: How can we derive x given R_d, R_{d^u}, and R_{d^p}? The answer is, we cannot. The structure of the equations is such that we cannot express x as a function of the known entities R_d, R_{d^u}, and R_{d^p}.[221] We can, however, give good lowerbound and upperbound estimates for x—meaning we can say that the actionable rate of confusion must be higher than some rate (the lowerbound) and lower than some other rate (the upperbound). We would like, of course, to give the greatest possible lowerbound and the least possible upperbound so we have a more precise estimate. That is, we would like to give the smallest possible interval that we know contains x. A good way of expressing the greatest lowerbound, call it $\underline{x}$, and the least upperbound, call it $\overline{x}$, so we have $x \in [\underline{x}, \overline{x}]$, is as follows[222]:

$$\underline{x} = max\{R_d - R_{d^u}, R_{d^p}\}$$
$$\overline{x} = min\{R_d - R_{d^u} + R_{d^p}, R_d\}$$

This means that we are taking the maximum of the quantities $R_3 + R_4$ and $R_1 + R_3$ to be the lowerbound estimate of x (which recall equals $R_1 + R_3 + R_4$). Depending on which of R_1 or R_4 is larger,[223] one or the other of these quantities will be the best (i.e., largest) lowerbound estimate. And we are taking the minimum of the two quantities $R_1 + R_3 + R_3 + R_4$ and $R_1 + R_2 + R_3 + R_4$ to be the upperbound estimate of x. Depending on which of R_2 or R_3 is larger,[224] one of these quantities will be the best (i.e., smallest) upperbound estimate.

---

the actionable confusion rate. Therefore, subtracting R_{d^u} from R_d yields an incorrect (and incorrectly low) estimate of actionable confusion. (Formally, $x \geq R_d - R_{d^u}$.)

221. Our inability to give a direct estimate of x is due to our choice of a between-subject survey design. *See supra* notes 136–40 and accompanying text. If we had chosen a within-subject design, with each respondent being exposed to d, d^u, and d^p (instead of to only one of them), then we could have given a direct estimate of x, by summing up (1) respondents who were confused in the d condition but not in the d^u condition (corresponding to the white balls and the dotted balls in Figure 9), and (2) respondents who were confused in all of the d, d^u, and d^p conditions (corresponding to the striped balls in Figure 9). However, as we explained in justifying our choice of a between-subject design, we are concerned that exposure to multiple conditions would affect a respondent's response to each condition, so the confusion rates obtained from a within-subject design would not be trustworthy.

222. Technically speaking, we will not prove that $[\underline{x}, \overline{x}]$ is the shortest possible interval estimate of x that can be computed based on R_d, R_{d^u}, R_{d^p}, but it is the shortest that can be computed based on straightforward, intuitively interpretable operations involving R_d, R_{d^u}, R_{d^p}. More complicated functions of those entities are unlikely ever to be adopted in litigation, so we do not consider them.

223. That is, depending on whether there are more people who are confused both by p_u alone and by p_p alone or people who are confused by (p_u, p_p) together but not p_u or p_p alone.

224. That is, depending on whether there are more people who are confused by p_u alone but not p_p alone or people who are confused by p_p alone but not p_u alone.

The legal interpretation is that the greatest lowerbound ($\underline{x}$) is our best defendant-friendly estimate, and the least upperbound ($\overline{x}$) is our best plaintiff-friendly estimate, of the actionable rate of confusion. The test given in the main text of Section II.B.1 (see the conclusions in Figure 10 and Table 1) is based on the greatest lowerbound estimate, so it's defendant-friendly.[225] This could be justified on the basis that the burden of showing likelihood of confusion falls on the plaintiff. But we can imagine arguments to the contrary, and a court might choose either the defendant-friendly or the plaintiff-friendly estimate or perhaps the entire range (which would be an interval estimate rather than a point estimate) as the rate of actionable confusion to be inferred from the surveys. Alternatively, it might be appropriate to apply the defendant-friendly standard to surveys done by plaintiffs and the plaintiff-friendly standard to surveys done by defendants, on the theory that the standard should compensate for the fact that the party administering the survey is doing as much as possible (within ethical and methodological bounds) to skew the results in its own favor.

Another nuance to be untangled pertains to the statement "cannot infer actionable likelihood of confusion" in our decision tree in Figure 10. It follows from the foregoing discussion that the two different paths leading to that conclusion (the "no" path and the "yes, no, no" path) represent *different* degrees of confidence in the conclusion. If the treatment survey reveals no appreciable confusion between p and d (the "no" path), then we know that the raw confusion rate (R_d) does not exceed the relevant threshold, so even the plaintiff-friendly standard would not allow us to infer actionable likelihood of confusion.[226] By contrast, if the treatment survey reveals appreciable confusion but we get negative results from the control surveys (the "yes, no, no" path), then the conclusion that we cannot infer actionable likelihood of confusion is based on the defendant-friendly standard. The adverse inference is stronger in the first instance.

B. *Deeper into the Weeds: The Real World*

We now analyze the realistic context where we do not assume away any confusion due to "inevitable" product features (denoted p_0)—meaning features, including non-trademark features, that are inevitably part of the product or its design or packaging. Likewise, we allow for respondents being confused by virtue of being asked to participate in a survey or simply by default. Note that, by definition of being inevitable, the inevitable features are always part of any product or control.

---

225. As discussed, the current approach of subtracting R_{d^u} from R_d is also a defendant-friendly standard, but our standard is more precise because it takes the maximum of two quantities rather than one of the two quantities as the greatest lowerbound. In other words, the current standard estimates a lowerbound that is not the best (i.e., greatest) lowerbound. What is more, it does not recognize that what it is estimating is a lowerbound (and not the best lowerbound, at that) and puts it forth as the best point estimate.

226. Of course, if R_d does not exceed the threshold, then $\overline{x}$ perforce does not exceed the threshold either; the minimum of two quantities is, by definition, no larger than any one of the two.

The allegedly infringed trade dress, p, now takes the form (p_0, p_u, p_p, p_3), the allegedly infringing trade dress is $d = (p_0, p_u, p_p, d_3)$, and the controls are $d^u = (p_0, p_u, d_2, d_3)$, $d^p = (p_0, d_1, p_p, d_3)$, and $d^0 = (p_0, d_1, d_2, d_3)$. The last control is specific to this context and was not present in the simplified context.

The aggregate rate of confusion is now represented by:

$$R_d = R_0 + R_1 + R_2 + R_3 + R_4$$

where

- R_0 represents those who are confused by p_0 alone (including those who are always confused)
- R_1 represents those who are confused both by (p_0, p_u) and by (p_0, p_{py}) but not by p_0 alone
- R_2 represents those who are confused by (p_0, p_u) but not by (p_0, p_p)
- R_3 represents those who are confused by (p_0, p_p) but not by (p_0, p_u)
- R_4 represents those who are confused by (p_0, p_u, p_p) together but not by (p_0, p_u) or (p_0, p_p) alone

As in the simplified version, this is a true partition.[227] And the notation is consistent with that in the simplified version except for the addition of R_0 and the inclusion of p_0 in R_1 through R_4. Note also that, by monotonicity, the sets R_2, R_3, and R_4 exclude those who are confused by p_0 alone (because of the "but not . . ." proviso in their definition).

The actionable rate of confusion is now given by

$$x = R_1 + R_3 + R_4$$[228]

227. Formally, $r_0 \cup r_1 \cup r_2 \cup r_3 \cup r_4 = r_d$ and $r_i \cap r_j = \emptyset \ \forall i \neq j, \ i, j \in \{0, 1, 2, 3, 4\}$.

228. Some readers might wonder why those confused by p_0, or at least some of them, are not included in the actionable confusion rate. An argument for including them would go as follows:

> If a person is confused by (p_0, p_p), then p_p is a sufficient cause of that person's confusion, so those who are confused by p_0 alone should count toward the actionable rate of confusion if they are also confused by (p_0, p_p). Failing to count these persons would be like failing to count the "striped balls" or R_1 (those confused by both p_u alone and p_p alone) in the simplified version. So, by failing to count them—that is, by subtracting R_{d^0} from R_{d^p} instead of using R_{d^p} in whole—we are making the exact same mistake that we accuse the current best practice of making.

This argument is incorrect. By monotonicity, *all* of those who are confused by p_0 alone are also confused by (p_0, p_p) (and, for that matter, also confused by (p_0, p_u), and (p_0, p_u, p_p)). So, the set of "those who are confused by p_0 alone and also confused by (p_0, p_p)" is the same as "those who are confused by p_0 alone." But if we were to include those who are confused by p_0 alone in the actionable rate of confusion, we would be counting *all* confusion as actionable confusion—doing away with controls altogether and throwing out the whole idea of causal tracing and the missing element. The analogy with the striped balls or R_1 in the simplified world is inapposite because, in that context, p_u and p_p are separate and separable elements such that a consumer could be confused

And the control surveys now yield

$$R_{d^u} = R_0 + R_1 + R_2$$
$$R_{d^p} = R_0 + R_1 + R_3$$
$$R_{d^0} = R_0$$

Thus, we have

$$R_d - R_{d^u} = R_3 + R_4$$

and

$$R_{d^p} - R_{d^0} = R_1 + R_3$$

So, as in the simplified version, our approach (see Figure 11) yields the lowerbound estimate

$$\underline{x} = max\{R_3 + R_4,\ R_1 + R_3\}$$

As before, this is a better estimate than that provided by the current best practice, which simply uses $R_3 + R_4$.

by one and not the other. Here, by contrast, p_0 is by definition *always* a part of the product, so we can never have p_p or p_u "alone."

C. NOTATION TABLE

Symbol	Meaning
p	Plaintiff's product or trademark (which could be a trade dress).
d	Defendant's product or trademark (which could be a trade dress).
$p = (p_u, p_p, p_3)$ or $p = (p_0, p_u, p_p, p_3)$	Plaintiff's product or trademark, broken into constituent elements, with the elements defined in the rows below.
$d = (p_u, p_p, d_3)$ or $d = (p_0, p_u, p_p, d_3)$	Defendant's product or trademark, broken into constituent elements, with the elements defined in the rows below. Plaintiff's and defendant's products share the same or similar elements p_0, p_u, and p_p but not p_3 or d_3.
p_0	Inevitable features of the plaintiff's and defendant's product—inevitable in the sense that one cannot design a comparable product that does not contain these features. For example, being an energy bar in disposable packaging.
p_u	Unprotectable features of the plaintiff's trademark that are common with or similar to defendant's trademark. For example, see-through wrapping in the Kind and Clif Mojo energy bars.
p_p	Protectable features of the plaintiff's trademark that are common with or similar to defendant's trademark. For example, the overall configuration of the Kind energy bar, which is similar to that of the Clif Mojo Bar.
p_3	Features of the plaintiff's trademark that are not common with or similar to defendant's trademark. For example, the Kind logo on the Kind energy bar.
d_3	Features of the defendant's trademark that are not common with or similar to plaintiff's trademark. For example, the Clif Mojo logo on the Clif Mojo energy bar.
$d^u = (p_u, d_2, d_3)$ or $d^u = (p_0, p_u, d_2, d_3)$	Control product that shares with plaintiff's trademark only the same or similar unprotectable features as defendant's product (p_u), such as see-through wrapping, as well as inevitable features (p_0) such as being an energy bar in disposable packaging. See Figure 6 and image Q15* on page 1305 for an example.
$d^p = (d_1, p_p, d_3)$ or $d^p = (p_0, d_1, p_p, d_3)$	Control product that shares with plaintiff's product only the same or similar protectable features as defendant's product, such as the Kind Bar's overall configuration. See image Q15* on page 1306 for an example.
$d^0 = (p_0, d_1, d_2, d_3)$	Control product that shares with plaintiff's product only its inevitable features and not any of its unprotectable or protectable trademark features. See Figure 5 and image Q15* on page 1307 for an example.