The Trademark Reporter®

INTA International Trademark Association

The Law Journal of the International Trademark Association

November–December, 2023 Vol. 113 No. 6

INTERNATIONAL TRADEMARK ASSOCIATION

675 Third Avenue, New York, NY 10017-5704
Telephone: +1 (212) 642-1700 email: wknox@inta.org

The views expressed in *The Trademark Reporter* (TMR) do not necessarily reflect those of the International Trademark Association (INTA). To fulfill its mission of delivering cutting-edge scholarship on trademarks, brands, and related intellectual property to its readers, the TMR sources content reflecting a diversity of viewpoints; the views expressed in any given article, commentary, or book review are those of the individual authors.

The TMR (ISSN 0041-056X) is published electronically six times a year by INTA, 675 Third Avenue, New York, NY 10017-5704 USA.

The Trademark Reporter®

(USPS 636-080)

Vol. 113 November–December, 2023 No. 6

TABLE OF CONTENTS

ARTICLES

COMMENTARY

This issue of *The Trademark Reporter* (TMR) should be cited as 113 TMR ___ (2023).

The Trademark Reporter®

A NEW ADDITION TO THE TRADEMARK LITIGATOR'S TOOL KIT: A NEUROSCIENTIFIC INDEX OF MARK SIMILARITY

By Mark Bartholomew,[*] Zhihao Zhang,[**] Ming Hsu,[***] Andrew Kayser,[****] and Femke van Horen[*****]

TABLE OF CONTENTS

[*] Professor of Law, School of Law, University at Buffalo.

[**] Assistant Professor of Business Administration, Darden School of Business, University of Virginia.

[***] Associate Professor of Business Administration, Haas School of Business, University of California, Berkeley.

[****] Professor of Neurology, School of Medicine, University of California, San Francisco.

[*****] Professor of Consumer Behavior, School of Business and Economics, Vrjie Universiteit Amsterdam.

I. INTRODUCTION

The linchpin for any claim of trademark infringement is an assessment of consumer confusion, but for a long time trademark practitioners and scholars have bemoaned the inherent difficulty of making such an assessment.[1] Judges cannot substitute their own views for those of the average consumer. Instead, trademark law demands the ability of the trier of fact "to think *through* the consumer and see the marketplace only as the consumer sees it."[2] Seeing through the eyes of the average consumer is easier said than done, however, as evident in the wide variety of judicial conceptions of consumer capabilities.[3] The judge must always worry that their sense of mark similarity or product proximity may be different from that of the relevant purchasing segment. Survey evidence presents a means for surfacing actual consumer perceptions and avoiding the trier of fact's own subjective experience, but its probative value is often discounted over fears of bias and inaccurate consumer self-reporting.

Enter neuroscience. Neuroscientific techniques promise a more unvarnished view of consumer perception, one that is not mediated through consumer self-reporting. Neuroscience is already influencing the law in a variety of areas, from tort law to the death penalty.[4] One difficulty for neuroscience, however, comes from the need to translate scientific understandings—typically generated from a group of research subjects participating under stable laboratory conditions—to a specific individual acting within the less constrained real world. In particular, most legal applications hinge upon the mental state of a particular person at a defined moment in time—e.g., what was the mental state of the killer at the moment of the crime?—something that goes beyond what can be provided by neuroscientific techniques, at least in the near term.

In contrast, trademark law's determination of infringement depends on the *aggregate* sense of consumers. Because likelihood of confusion—the issue at the center of any claim of trademark

---

1 *E.g.*, Robert G. Bone, *Notice Failure and Defenses in Trademark Law*, 96 B.U. L. Rev. 1245, 1255-56 (2016) (criticizing the vagueness and variable application of the likelihood of confusion test); William E. Gallagher & Ronald C. Goodstein, *Inference Versus Speculation in Trademark Litigation: Abandoning the Fiction of the Vulcan Mind Meld*, 94 TMR 1229, 1231 (2004) (suggesting that "trademark law practitioners cannot safely assume that we are fairly representative of the class of relevant consumers").

2 Barton Beebe, *Search and Persuasion in Trademark Law*, 103 Mich. L. Rev. 2020, 2022 (2005).

3 Laura A. Heymann, *Trademark Law and Consumer Constraints*, 64 Ariz. L. Rev. 339, 340-41 (2022).

4 Roper v. Simmons, 543 U.S. 551, 569-78 (2005) (citing neuroscientific evidence in barring the death penalty for crimes committed by juvenile offenders); Teneille R. Brown, *Minding Accidents*, 94 U. Colo. L. Rev. 89, 119 (2023) (using neuroscientific findings to argue for changes to the law of negligence).

infringement—asks about the overall perceptions of the relevant consuming pool, it avoids the difficulty of extrapolating from aggregate data to a specific instance (often referred as the "group to individual" or "G2i" problem[5]), by probing a more enduring mental representation, and can thus more immediately benefit from neuroscientific evidence in a way that other legal questions (e.g., the *mens rea* of a particular criminal defendant at a particular moment) cannot.[6]

This is not to say that brain scans can somehow substitute for the entire likelihood of confusion analysis. Just as survey evidence is only one potential part of a larger holistic assessment of consumer confusion, neuroscientific data on consumer perception would serve as an additional resource, not a replacement. Ultimately, where to set the line between infringing and non-infringing conduct is up to lawmakers, not scientists, as the data itself provides factual information but not normative judgments. Still, additional insight into consumer thought is bound to be helpful: "As long as trademark purports to be guided by consumer reactions, it can only benefit from a better understanding of those reactions."[7]

This article makes the case for neuroscience's value in improving assessments of mark similarity, and thereby consumer confusion. Part II describes how likelihood of confusion is determined in trademark law, chronicling long-standing frustrations with confusion's assessment and current doctrinal shortfalls. It also takes some time to examine the shortcomings with survey evidence of confusion. This review sets the stage for Part III, which offers our proof of concept.[8] We describe an experiment we conducted to

---

[5] David L. Faigman et al., *G2i Knowledge Brief: A Knowledge Brief of the MacArthur Foundation Research Network on Law and Neuroscience*, MacArthur Foundation Research Network on Law and Neuroscience (2016) ("Even the best science—science characterized by rich data collected from multiple experimental subjects or events and over multiple trials or experiments—frequently can tell us little, if anything at all, about the individual case. …Scientists typically don't attempt to infer from group or population-based data (or 'G') to a particular individual (or 'I')."); David L. Faigman, John Monahan & Christopher Slobogin, *Group to Individual (G2i) Inference in Scientific Expert Testimony,* 81 U. Chi. L. Rev. 417, 420 (2014) ("This gap between conventional scientific practice and ordinary trial practice involves the challenge of reasoning from group data to decisions about individuals (an analytical process that we designate as 'G2i').").

[6] For an in-depth discussion of neuroscience's potential to shed light not only on trademark infringement, but other mental states at issue in trademark, patent, and copyright law, see Mark Bartholomew, Intellectual Property and the Brain: How Neuroscience Will Reshape Legal Protection for Creations of the Mind (2022).

[7] Rebecca Tushnet, What's the Harm of Trademark Infringement?, 49 Akron L. Rev. 627, 646 (2016).

[8] Our neuroscience paper was published in a peer-reviewed general science journal, *Science Advances*. *See* Zhihao Zhang, Maxwell Good, Vera Kulikov, Femke van Horen, Mark Bartholomew, Andrew S. Kayser & Ming Hsu, *From Scanner to Court: A Neuroscientifically Informed "Reasonable Person" Test of Trademark Infringement*, 9 Sci. Advances eabo1095 (2023), https://www.science.org/doi/ 10.1126/sciadv.abo1095.

construct an index of neural responses to visual similarity. This index relies on a well-established neuroscientific phenomenon in which the neural response declines upon repeated presentation of the same stimulus, thereby permitting the presentation of different marks and a corresponding calculation of their perceived similarity. Our research demonstrates that mark similarity is capable of quantification, and in a way that avoids some of the key problems plaguing survey research. Part IV offers thoughts on how neuroscientific measurements like ours can be deployed to improve not only the likelihood of confusion analysis, but other areas of trademark law that also depend on understanding of aggregate consumer sentiment like secondary meaning and dilution.

II. PROBLEMS IN ESTIMATING CONFUSION

As readers of this journal will no doubt be aware, U.S. courts estimate the likelihood of consumer confusion through a multi-factor test. Though the specific number and description of these factors vary according to jurisdiction, every federal circuit evaluates the same basic factors to determine confusion:

- Similarity of the plaintiff's and defendant's marks
- Strength of the plaintiff's mark
- Intent of the defendant
- Purchaser sophistication
- Presence of actual confusion
- Relatedness of the goods or services at issue

A read of these factors reveals that trademark law's confusion analysis hinges on proxies for consumer perception rather than direct evidence of that perception. Only the actual confusion factor permits direct evidence of consumer thought to enter the analysis, and proof of actual confusion is not necessary to demonstrate likelihood of confusion—the rest rely on circumstantial evidence that a judge or jury will decide how to weigh.

The result has been a certain amount of judicial angst when it comes to determining trademark infringement. In 1948, Judge Jerome Frank famously referred to application of the factors as a "shaky kind of guess." [9] He even undertook his own questioning of a randomly chosen group of "adolescent girls and their mothers and sisters" to decide whether MISS SEVENTEEN girdles would appear confusingly similar to SEVENTEEN magazine.[10] A great deal of more modern commentary explores how legal adjudications of confusion under the above factors are likely to differ from actual

---

[9] Triangle Publ'ns, Inc. v. Rohrlich, 167 F.2d 969, 976 (2d Cir. 1948).

[10] *Id.*

consumer sentiment.[11] Despite the original intent of a structured and principled approach to weighing the different factors, judges can base their judgments on their own intuition and informed gut feeling[12] and may evaluate factors in a way that is "consistent with the outcome they favor on other grounds" rather than independently.[13] One court of appeals warned against giving outsized weight to any one confusion factor (like mark similarity) because this approach could be a mechanism "where the subjective impressions of a particular judge are weighed at the expense of other relevant evidence."[14]

Given this concern with judicial subjectivity, it is no wonder that consumer surveys are often submitted into evidence and can take on great importance in deciding a trademark infringement claim. Although there is some disagreement as to their overall significance,[15] there is little doubt that surveys can be critical, and sometimes determinative, in trademark litigation.[16] Survey evidence is important not just in adjudicated cases, but in evaluating the strength of infringement claims in pretrial litigation.[17]

---

[11] Martin Senftleben & Femke van Horen, *The Siren Song of the Subtle Copycat—Revisiting Trademark Law with Insights from Consumer Research*, 111 TMR 739, 741 (2021) (maintaining that legally modest amounts of similarity are more likely to influence consumers' purchasing decisions than blatant forms of similarity); Jeanne C. Fromer & Mark A. Lemley, *The Audience in Intellectual Property Infringement*, 112 Mich. L. Rev. 1251, 1260 (2014) (specifying ways the jury's perspective is likely to differ from that of actual consumers in trademark infringement cases).

[12] *See* D.J.G. Visser, *Beslissen in IE-zaken* [Deciding IP Cases], 31 NJB 1918 (2008) (presenting results from survey of seventeen judges on how they decide intellectual property disputes).

[13] Shari Seidman Diamond & David J. Franklyn, *Trademark Surveys: An Undulating Path*, 92 Tex. L. Rev. 2029, 2043 (2013).

[14] Jada Toys, Inc. v. Mattel, Inc., 518 F.3d 628, 632-34 (9th Cir. 2008).

[15] Irina D. Manta, *In Search of Validity: A New Model for the Content and Procedural Treatment of Trademark Infringement Surveys*, 24 Cardozo Arts & Ent. L.J. 1027, 1029 (2007) ("[W]ithout survey evidence it is generally almost impossible to prove trademark infringement."); Kevin Blum, Ariel Fox, Christina J. Hayes & James (Hanjun) Xu, *Consistency of Confusion? A Fifteen-Year Revisiting of Barton Beebe's Empirical Analysis of Multifactor Tests for Trademark Infringement*, 2010 Stan. Tech. L. Rev. 3, 30 ("[S]urvey data is less frequently employed than one might expect given the conventional wisdom that survey evidence is routinely employed to prove a likelihood of confusion.").

[16] Blum et al., *supra* note 15, at 30 (noting that "91.7% of the opinions crediting . . . survey evidence also found in favor of the party presenting the survey"); *see also* Eric D. DeRosia, *Fixing* Ever-Ready*: Repairing and Standardizing the Traditional Survey Measure of Consumer Confusion*, 53 Ga. L. Rev. 613, 617 (2019) (stating that confusion "surveys frequently play an important role in pretrial negotiations and at trial" (footnote omitted)).

[17] Diamond & Franklyn, *supra* note 13, at 2061-62. Other means besides surveys for assessing confusion have been offered over the years. *See, e.g.*, R. Bradlee Boal, *Techniques for Ascertaining Likelihood of Confusion and the Meaning of Advertising Communications*, 73 TMR 405, 407-408 (1983) (describing "in store" coupon test); Jean-Noël Kapferer, *Brand Confusion: Empirical Study of a Legal Concept*, 12 Psych. & Mktg.

Nevertheless, despite their common usage, there is also widespread skepticism about the value of confusion surveys. The Second Circuit cautions that surveys are "not immune to manipulation."[18] Judge Posner of the Seventh Circuit warned of the potential for surveys to be steered in a desired direction, referencing "the survey researcher's black arts."[19] A concern in the litigation between Jack Daniel's and VIP Products (maker of "Bad Spaniels" dog chew toys) before the U.S. Supreme Court in 2023 was the perceived "precarity of consumer surveys."[20] As part of its case for confusion, Jack Daniel's touted a survey reporting that twenty-nine percent of those shown photographs of the "Bad Spaniels" toy identified Jack Daniel's as making, sponsoring, or approving it.[21] But amicus briefs filed with the Supreme Court contended that survey evidence, particularly in a parody case, should not be considered reliable enough to force a defendant to endure a lengthy and expensive legal investigation of consumer confusion.[22] These concerns seemingly resonated with Justices Sotomayor and Alito, who authored a concurrence warning of the "risk in giving uncritical or undue weight to surveys."[23] This skepticism is based on three chief criticisms of confusion survey evidence: (1) it can plant the idea of confusion in respondents' heads, leading to artificially high readings of confusion; (2) it may include subtle manipulations of wording that steer results; and (3) it can fail to capture the actual nuances of consumer perception of similarity.

---

551 (1995) (testing for confusion using a tachistoscopic experiment); Takuya Satomura, Michel Wedel & Rik Pieters, *Copy Alert: A Method and Metric to Detect Visual Copycat Brands*, 51 J. Mktg. Rsch. 1 (2014) (proposing a three-part method for assessing confusion from visual similarity). These have their own problems, however, and none of them has gained traction in trademark litigation.

18 Nabisco, Inc. v. PF Brands, Inc., 191 F.3d 208, 224 (2d Cir. 1999).

19 Indianapolis Colts, Inc. v. Metro. Baltimore Football Club Ltd. P'ship, 34 F.3d 410, 416 (7th Cir. 1994).

20 *See* Brief of Amici Curiae 30 Trademark Law Professors in Support of Respondent at 15, Jack Daniel's Props., Inc. v. VIP Prods. LLC, 143 S. Ct. 1578 (2023) (No. 22-148), http://www.supremecourt.gov/DocketPDF/22/22-148/255429/20230223152033068_22-148%20Trademark%20Law%20Professors%20Amicus%20Brief%20TO%20FILE.pdf.

21 VIP Prods., LLC v. Jack Daniel's Props., Inc., 291 F. Supp. 3d 891, 907 (D. Ariz. 2018).

22 *See* Brief of Amici Curiae 30 Trademark Law Professors in Support of Respondent, *supra* note 20, at 17-19; Brief of Amici Curiae First Amendment Professors in Support of Respondent at 17-18, Jack Daniel's Props., Inc. v. VIP Prods. LLC, 143 S. Ct. 1578 (2023) (No. 22-148), https://www.supremecourt.gov/DocketPDF/22/22-148/255392/20230223130058120_43246%20pdf%20Tushnet%20br.pdf; Brief of Amicus Curiae Foundation for Individual Rights and Expression in Support of Respondent at 9, Jack Daniel's Props., Inc. v. VIP Prods. LLC, 143 S. Ct. 1578 (2023) (No. 22-148), https://www.supremecourt.gov/DocketPDF/22/22-148/255433/20230223152247858_22-148%20Amicus%20Brief.pdf ("[I]f consumers wrongly assume that satirical use of another's mark requires the owner's permission—or if the mark owner manipulates a survey to show that—even an obvious parody can succumb to a finding of affiliation confusion.").

23 Jack Daniel's Props., Inc. v. VIP Prods. LLC, 143 S. Ct. 1578, 1593 (2023) (Sotomayor, J., concurring).

Just the effort to poll respondents about confusion can prompt them to conclude that confusion exists. In the dominant *Ever-Ready*[24] survey format, respondents are shown the junior brand and asked: (1) "Who do you think puts out this brand?," (2) "What makes you think so?," and (3) "Name any other products put out by the same concern which puts out this brand." Responses that name the senior user are evidence of confusion. But a criticism of this format is that it may plant a seed in respondents' heads, nudging them to think of possible connections to another's mark, whereas, in a normal shopping trip without the prompt, they would not have made a connection to anyone at all.[25]

The other prominent method for surveying trademark confusion is the *Squirt*[26] method. It asks, "Do you think [the senior mark] and [the junior mark] are put out by the same company or by different companies?" This test has the advantage of surfacing potential confusion in situations where consumers were not previously aware of the senior mark. But it also has the potential to skew results in favor of finding confusion. Because the *Squirt* method explicitly asks subjects to consider the association between the two marks, it threatens to cause the subjects to identify an association that might not exist in a typical purchasing transaction.[27] In the words of the Tenth Circuit, by pairing the junior and senior marks together, the *Squirt* method can "suggest[] the very answer most helpful to [the senior user's] cause."[28]

Separate from the issue of artificially disposing research subjects to be attuned to confusion, survey evidence is also plagued by more general issues over question wording. Researchers document how even subtle differences in the wording of the *Ever-Ready* questions can influence survey results either above or below the legal threshold for infringement.[29] Due to these "demand effects," respondents may use cues provided by the survey procedures or questions, causing them to modify their answers in a way that aligns with what they perceive as the goals or expectations of the survey. Leading questions, such as "Do you believe that this

---

[24] Union Carbide Corp. v. Ever-Ready, Inc., 531 F.2d 366, 385-88 (7th Cir. 1976).

[25] Itamar Simonson, *The Effect of Survey Method on Likelihood of Confusion Estimates: Conceptual Analysis and Empirical Test*, 83 TMR 364, 369 (1993); *see also* DeRosia, *supra* note 16, at 620 (contending that there are actually several variants of the *Ever-Ready* survey that have been accepted by the courts and these variants can produce significantly different responses in survey respondents); E. Deborah Jay, *He Who Steals My Good Name: Likelihood of Confusion Surveys in TTAB Procedures*, 104 TMR 1141, 1159 (2014) ("Eveready surveys are more effective at proving that confusion is likely than at proving it is unlikely.")

[26] Squirtco v. Seven-Up Co., 628 F.2d 1086, 1089 n.4, 1091 (8th Cir. 1980).

[27] Simonson, *supra*, at 371.

[28] Water Pik, Inc. v. Med-Systems, Inc., 726 F.3d 1136, 1148 (10th Cir. 2013).

[29] DeRosia, *supra* note 16, at 620.

restaurant is connected with or related to any other restaurants?," clearly suggest the expected answer.[30] Such demand effects can significantly bias survey findings.

Finally, beyond a generalized concern with the potential for manipulation, surveys are attacked for failing to capture the nuances of human perception. For example, a group of researchers recently complained that survey evidence fails to take into account the relative certainty of a respondent's judgment of similarity.[31] As currently used, the researchers argued, surveys are too blunt of a tool to deserve much credence in determining the outcome of a claim of infringement. By too often forcing the subject into a binary choice—is there an association between the two marks at issue or not?—most trademark surveys neglect to take into account the relative strength of consumer perceptions of mark similarity, something that is likely to impact true confusion on the ground.

Common objections to potentially misleading survey techniques have not coalesced into something approaching a recognizable template for judges and litigants.[32] One need only look at past issues of this journal to see that, despite numerous attempts to document specific shortcomings in surveys, surveys with those shortcomings continue to be deployed in trademark litigation.[33] Though the risk of bias in confusion surveys is widely known by judges and litigators, it is difficult to actually demonstrate that bias, except for outliers, in such a way as to get a survey deemed inadmissible under the *Daubert* threshold for expert testimony.[34] Put another way, no gold standard exists for demonstrating flaws in trademark surveys. The

---

30 Itamar Simonson & Ran Kivetz, *Demand Effects in Likelihood of Confusion Surveys: The Importance of Marketplace Conditions*, *in* Trademark and Deceptive Advertising Surveys 243 (Shari Seidman Diamond & Jerre B. Swann eds., 2d ed. 2022).

31 Barton Beebe, Roy Germano, Christopher Jon Sprigman & Joel H. Steckel, *Consumer Uncertainty in Trademark Law: An Experimental Investigation*, 72 Emory L.J. 489 *passim* (2023).

32 There are indeed works that provide guidance for performing trademark surveys. *See, e.g.*, Trademark and Deceptive Advertising Surveys, *supra* note 30. Yet despite repeated attempts to set a standard for such surveys, many questions as to appropriate survey design and consideration remain unsettled. *See* Manta, *supra* note 15, at 1029 ("[T]he standards governing the treatment of surveys in trademark infringement cases are vague and unclear, which leads to confusion in the legal community and leaves trademark owners unable to ensure the protection of their intellectual property.").

33 *E.g.*, John P. Liefeld, *How Surveys Overestimate the Likelihood of Confusion*, 93 TMR 939, 939-40 (2003) (noting the increase in court acceptance of surveys in trademark cases); Michael Rappeport, *A Replication Problem in Survey Design, Including a Critique of the Decision in* Thoip v. Disney, 100 TMR 1360, 1363 (2010) (observing that both the USPTO and Trademark Trial and Appeal Board commonly accept and rely on survey evidence).

34 Artemio Rivera, *Testing the Admissibility of Trademark Surveys After* Daubert, 84 J. Pat. & Trademark Off. Soc'y 661, 663 (2002) ("In spite of *Daubert*, the conventional wisdom in trademark litigation remains that the existence of flaws in the design or implementation of a survey does not raise an admissibility issue, and instead must only be considered by the fact finder in weighing evidence.").

result can be rival surveys that testify to wildly different rates of confusion, with the trier of fact left to throw up their hands and turn to the other likelihood of confusion factors.

In summary, despite a growing judicial acceptance of confusion surveys in trademark cases, concerns remain as to their validity and reliability.[35] The nature of surveys requires that respondents be confronted with language prompts that can cause them to be more vigilant or aware of the potential for confusion than in the everyday context that trademark law is meant to assess. Survey design includes subtle language choices that can influence outcomes, intentionally or not, in particular directions. The predominant confusion survey formats do not capture the nuances of consumer perception, including the strength of those perceptions. In addition, because respondents do not always have perfect insight into their own perceptions or may be inclined to tell researchers what they think they want to hear, surveys can be seen as flawed instruments regardless of question wording.

But are neuroscientific assessments of consumer confusion any better? For neuroscientific evidence to be valuable, it has to offer something surveys do not. We believe that brain imaging can offer evidence that avoids some of the problems found in survey data. By eliciting measures of confusion through a passive viewing task without the need to ask any (leading) questions, neuroscience can generate different and potentially more accurate assessments of confusion. Neuroscientific evidence of confusion—by recording responses to particular stimuli without asking questions about those stimuli—avoids one source of potential bias that has generated skepticism about survey evidence. Because it does not involve self-reporting, neuroscientific evidence eliminates the issues of mediation by research subjects that can skew reports of confusion. Neuroscientific measurements can also capture the degree of perceived mark similarity, improving on the simple "yes" or "no" measurements of most trademark confusion surveys.

Neuroscience is not a panacea—as we discuss, this kind of evidence of confusion can present its own issues. Another common criticism of survey evidence is that it fails to capture real world shopping conditions, and the same complaint can be lodged even more strongly against experiments that require representative consumers to lie still in a machine that measures their brain

---

[35] *See* Kraft Foods Grp. Brands LLC v. Cracker Barrel Old Country Store, Inc., 735, 741 (7th Cir. 2013) (explaining that confusion surveys "conducted by party-hired expert witnesses are prone to bias"); J. Thomas McCarthy, McCarthy on Trademarks and Unfair Competition § 32:196.73 (5th ed. 2023) (referring to "lingering judicial skepticism about survey evidence"); Beebe et al., *supra* note 31, at 546 ("It has now been a century since courts first began to consider trademark survey evidence, and through the course of that century, the quality and utility of survey evidence has substantially deteriorated, leaving judges understandably wary of it.").

activity.[36] Despite these limitations, however, this new kind of evidence offers a means for testing survey evidence for bias. Trademark disputes often feature competing surveys, with each side submitting evidence testifying to the presence or absence of confusion. Neuroscientific assessment of visual similarity might reveal the presence of bias in one of these submitted surveys. It can also serve as a useful tool before litigation commences to get a read on the likelihood of confusion from the defendant's activities. In the next part, we offer our proof of concept.

III. PROOF OF CONCEPT: A NEUROSCIENTIFIC READING OF CONFUSION

A. Similarity as a Driver (and Proxy) of Confusion

Is it possible to design a brain-based method to detect the presence of consumer confusion? The short answer is no, as neither the law nor cognitive neuroscience currently allows for this possibility. Legally speaking, "confusion" is too indeterminate. Legal scholars have criticized the ambiguity of what trademark confusion means and how it should be operationalized,[37] such that, in practice, credit is given by many judges to "any possible way that consumers might be confused."[38] From a neuroscientific standpoint, confusion is likely not a unitary mental state. Depending on the specific context, it may consist of a range of components, including processes related to sensory inputs (e.g., whether one is able to distinguish between the appearances of two or more products), memory (e.g., feelings of familiarity driven by activation of past experiences), reasoning (e.g., inferences about what company puts out a certain brand), emotion (e.g., reactions evoked by information ambiguity and processing ease), and decision-making (e.g., purchase of a product inconsistent with the original goal).[39] The complexity of

---

[36] At the same time, it is important to note that fMRI is by no means the only method available to cognitive neuroscientists. Indeed, there has been significant progress in developing wearable devices for neuroscientific measurement, potentially even allowing for data collection while consumers are actually shopping in the future. *See, e.g.*, Tim R. Mullen, Christian A.E. Kothe, Yu Mike Chi, Alejandro Ojeda, Trevor Kerth, Scott Makeig, Tzyy-Ping Jung & Gert Cauwenberghs, *Real-time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG*, 62 IEEE Transactions on Biomedical Eng'g 2553 (2015).

[37] 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, 1338 (2012) ("The Lanham Act does not define 'confusion,' and the likelihood of confusion test itself does nothing to clarify the meaning of the term."); Daryl Lim, *Trademark Confusion Revealed: An Empirical Analysis*, 71 Am. U. L Rev. 1285, 1287 (2022) ("[W]hat constitutes 'confusion' remains highly subjective and difficult to evaluate.").

[38] Bone, *supra* note 37, at 1338.

[39] *See, e.g.*, Vincent-Wayne Mitchell, Gianfranco Walsh & Mo Yamin, *Towards a Conceptual Model of Consumer Confusion*, 32 Advances Consumer Rsch. 143 (2005)

confusion will likely lead to highly variable and distributed brain activity patterns, a hypothesis that, to the best of our knowledge, has not been systematically studied.

Hence, neuroscience's potential for improving the likelihood of confusion analysis needs to be more modest. In our proof of concept,[40] we opted to examine perceived visual similarity as the focal mental state to be measured with neuroimaging. Not only is it widely recognized as a key driver of consumer confusion,[41] empirical studies of legal decision-making have shown that, of all the likelihood of confusion factors, assessments of visual similarity typically exert the greatest weight on the court's judgment.[42] Importantly, focusing on visual similarity affords the advantage of leveraging a rich literature on human visual processing in cognitive neuroscience. Arguably the best understood set of mental processes in modern neuroscience, visual perception has well-delineated underlying brain regions that support different components of the process, along with an established set of methodological tools for mapping brain activities to subjective mental states. More specifically, the representation of visual objects (such as consumer products) can be reliably and consistently measured in known parts of the human brain using neuroimaging.[43] While neuroscience cannot determine whether there is "confusion," it can provide empirical evidence of perceived visual similarity.

B. Tools for Measuring Perceived Visual Similarity Through Brain Activities

For those unfamiliar with neuroscience, a brief primer may be useful. A variety of techniques have been used in the long history of research on how visual information is processed in the brain. In our study, we used functional magnetic resonance imaging ("fMRI"), a widely used technique in cognitive neuroscience, to measure fine-grained brain activity patterns in healthy research volunteers, a

---

(discussing variables that can affect confusion); Markus Schweizer, Alexander J. Kotouc & Tillmann Wagner, *Scale Development for Consumer Confusion*, 33 Advances Consumer Rsch. 184 (2006) (same).

40 Zhang et al., *supra* note 8.

41 Barbara Loken, Ivan Ross & Ronald L. Hinkle, *Consumer "Confusion" of Origin and Brand Similarity Perceptions*, 5 J. Pub. Pol'y & Mktg. 195, 195 (1986) ("[S]imilarity in physical appearance of two brands (e.g., a store brand and a national brand) is significantly related to consumer perceptions of a common business origin between them.").

42 Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 Calif. L. Rev. 1581, 1600 (2006); Lim, *supra* note 37, at 1329.

43 *See* Leila Reddy & Nancy Kanwisher, *Coding of Visual Objects in the Ventral Stream*, 16 Current Op. Neurobiology 408 (2006) (using neuroimaging to evaluate how the brain represents visual objects).

population from which samples representative of typical consumers of common products can be drawn.

fMRI uses a strong magnetic field and radio waves to create pictures of the brain. During a given mental activity, brain regions that are more strongly engaged require more oxygen. The presence of oxygenated hemoglobin—the protein that carries oxygen in the blood—changes the magnetic properties of the blood, which in turn affects the local magnetic field in the vicinity of the active brain regions. Such small changes of the magnetic field, called blood oxygen level-dependent ("BOLD") signals, can be recorded by the fMRI scanner. By comparing the BOLD signal of a certain brain region during specific tasks or stimulus presentations to control conditions, researchers can identify not only which brain regions are involved in processing certain information, such as visual input, but also how such information is represented in the brain.

Here, our goal is to develop a tool that uses brain activity to more directly measure the perceived visual similarity between two trademarks. To achieve this goal, we leverage two specific sets of knowledge that have both been well established in cognitive neuroscience. First, the processing of visual objects, of which trademarks and trade dresses are specific instances, has been shown to proceed along a distinct pathway in the brain. This so-called "ventral pathway" starts from the back of the brain (the primary visual cortex), extracting basic visual features such as edges, orientations, and contrasts, to the sides of the brain (parts of the temporal lobe), where a rich, holistic representation of the objects and scenes is formed.[44]

Second, the perceived similarity between two stimuli can be measured directly, using a known property of brain responses called "repetition suppression."[45] Repetition suppression ("RS") is a phenomenon whereby the brain's response to a repeated stimulus decreases over time. For example, if a picture of a face is presented multiple times, the response in the brain regions responsible for processing that face gradually decreases. In the visual system, RS is a highly regular phenomenon that generalizes across people, and even across different mammalian species, such that robust measures can be derived with a relatively small sample of participants.

---

[44] *See* James J. DiCarlo, Davide Zoccolan & Nicole C. Rust, *How Does the Brain Solve Visual Object Recognition?*, 73 Neuron 415 (2012).

[45] *See* Helen C. Barron, Mona M. Garvert & Timothy E.J. Behrens, *Repetition Suppression: A Means to Index Neural Representations Using BOLD?*, 371 Phil. Transactions Royal Soc'y B: Biological Sci. 20150344 (2016).

Repetition suppression is thought to occur because the brain becomes less sensitive to stimuli that are repeated. Although the underlying neurobiological mechanism remains debated, it is believed that one important consequence of RS is to allow the brain to filter out irrelevant or unchanging information and focus on processing new and important information. To measure similarity between stimuli, we can therefore examine the extent of neural response reduction when similar stimuli are repeated. The idea is that if two stimuli are highly similar, the brain's response to the second stimulus will be more suppressed compared with a less similar stimulus (Figure I).

Figure I. Predicted brain response to the brand stimuli based on repetition suppression.

C. Our Experiment

The goal of our experiment was to test whether the degree of suppression observed in the object-sensitive area of the brain could be used to construct a brain-based index of perceived visual similarity.[46] To create a realistic simulation of legal cases, we chose two scenarios involving potential trademark infringement in the United States. Specifically, we picked two popular products, REESE'S Peanut Butter Cups and OXICLEAN laundry detergent.

---

[46] *See* Zhang et al., *supra* note 8 (including a detailed description of all methodology referenced in the text).

REESE'S candy was selected because of its involvement in a previous lawsuit against the import of a British candy called TOFFEE CRISP.[47] We included OXICLEAN detergent to introduce visual variations, such as color, and to evaluate a non-food item.

For each category (candy and cleaning product), we created a set of comparison products with varying visual similarities, determined through pretests. The inclusion of these comparison products ensured that instances covering a wide range of similarity, from highly dissimilar to highly similar, were included in our experiment, which would help us assess the effectiveness of our proposed neural index in distinguishing between different levels of similarity. Some stimuli, like TOFFEE CRISP candy and TIDE detergent, were based on real products, while others, such as "Pieces" peanut butter cups and "Breeze" detergent, were fictional or not sold in the United States. Additionally, we included two real product variants, REESE'S STICKS (a brand extension of REESE'S Peanut Butter Cups) and an international version of OXICLEAN detergent, which were intended to be highly similar, but not identical to, the actual REESE'S Peanut Butter Cups and OXICLEAN detergent, respectively. Hereafter we refer to REESE'S and OXICLEAN as the "reference product" for their respective categories, while other products are referred to as "competitor products."

During the experiment, participants were shown rapid presentations of product images in one of three different viewing angles (Figure II). Importantly, to induce RS, the competitor products and the reference product were grouped together to create pairs specific to each product category. In these pairs, a competitor product was followed by the reference product after a short interval. Additionally, pairs consisting of two consecutive presentations of the reference product were included, thereby anchoring one end of the similarity continuum with an identical stimulus pair.[48]

---

47 First Amended Complaint, Hershey Co. v. Posh Nosh Imports (USA) Inc., No. 2:14-cv-04028 (C.D. Cal. July 23, 2014), ECF No. 20, available at https://www.courtlistener.com/docket/4151437/the-hershey-company-v-posh-nosh-imports-usa-inc/.

48 To reduce the likelihood that participants identified such patterns, "spacer trials" were introduced, such that single presentations of the competitor products were randomly interspersed.

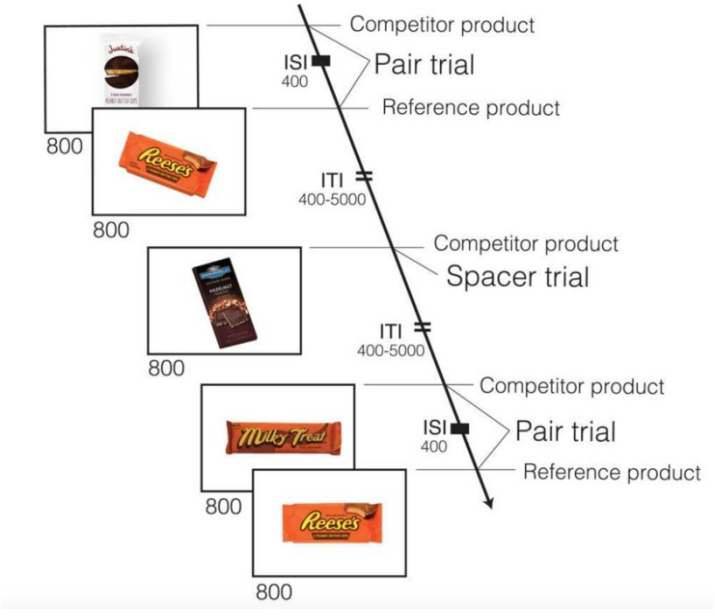


Figure II. Example trial timeline of the fMRI experiment.[49]

To avoid influencing participants' reactions, we did not inform them about the background or purpose of the study. Instead, they were given an unrelated task of identifying occasional inverted images. Specifically, they were instructed to press a button whenever they saw an inverted image, which appeared approximately once every few trials in a pseudorandom order. This so-called "cover" task served the purpose of keeping the participants engaged with the experiment by encouraging close attention to the visual stimuli.[50] In conjunction with the task, we also identified specific areas in the brain that were sensitive to objects. Together, this task design and this regional brain localization enabled us to extract the neural responses in the object-sensitive cortex for the stimulus pairs in the main task. These responses defined our neural similarity index. Specifically, this index was scaled to cover an interval between 0 and 1, corresponding to highest and lowest levels of similarity, respectively. For the upper end of the scale, we used the RS effect elicited by consecutive presentations of the same

[49] All product images were presented for 800 milliseconds in each trial. ITI, or the inter-trial interval, stands for the time interval between consecutive trials in the experiment, and was set to be 400 milliseconds. ISI, or the inter-stimulus interval, represents the time interval between the pair of products in the case of a pair trial, and ranged from 400 to 5000 milliseconds.

[50] Data from trials with inverted images were excluded from our analysis, and therefore would not affect the results on the neural index of similarity.

reference product because the reference product is most similar to itself. For the lower end of the scale, we used the competitor product with the weakest RS effect. As such, a similarity index between 0 and 1, based on the fMRI data reflecting the degree of RS, could then be calculated for each product.

D. Benefits of Our Neuroscientific Approach Over Surveys

The approach we employed offers several key advantages over survey-based methods, reducing potential biases that can be introduced through explicit self-report of perceived similarity. These advantages stem from the use of a direct measure of the brain activities associated with visual perception, along with a passive viewing paradigm in which participants are not actively asked to make similarity judgments (Figure III).

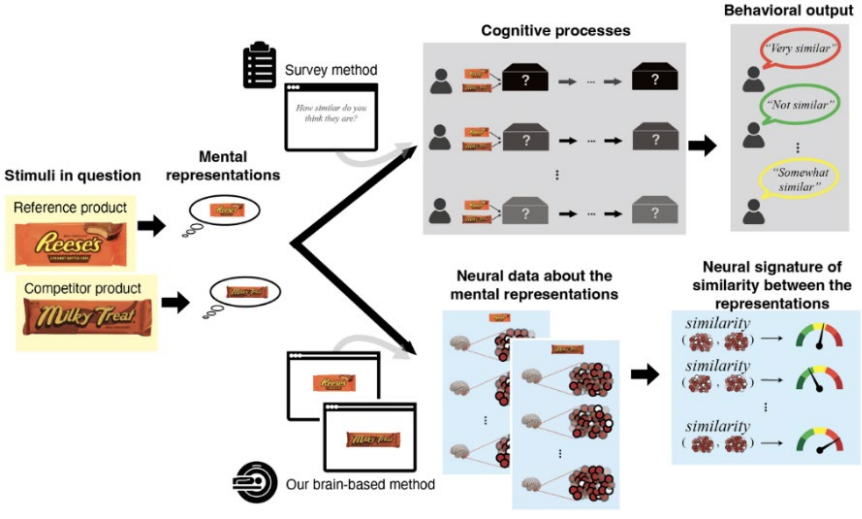


Figure III. Our brain-based measure (bottom) bypasses the complex cognitive processes involved in responding to a survey (top).[51]

---

[51] The stimuli in question (e.g., the package designs of two products) generate mental representations, which can be probed by different methods based on distinct assumptions. Survey methods (top) are based on the respondents' own assessments about the relationship between the mental representations and therefore recruit a series of additional cognitive processes that are not fully understood. This approach rests on the assumption that these processes are effectively shielded from biases and undue influences. Our neuroscientific approach using RS of the BOLD signals measured by fMRI (bottom) bypasses these processes, thereby providing a readout of the similarity between stimuli based on the neural correlates of their mental representations. This approach relies on the assumption, among others, of a reliable mapping between such representations and their neural correlates. The size of the RS effect observed in the

The first advantage is that by eliciting neural responses from participants without requiring any verbal responses, this paradigm removes the possibility of leading questions. In other words, participants are not actively engaging in (instructed) cognitive reasoning or decision-making while viewing the stimuli, allowing us to capture more direct and unbiased neural responses stemming from visual perception of the brands of interest. This design reduces the potential for biases introduced through arbitrary instructions related to different aspects of the judgment process—for instance, which features merit attention, and what counts as similar. Because the task does not address explicit comparisons at all, it also avoids the ambiguity introduced when comparisons are prompted but decision criteria are not explicitly provided.

The second advantage of passive viewing is that it enables blinding of the participants (and potentially administrators of the experiment) to not only the purpose of the study, but also the need to rate or compare the products and their associated trademarks at all. This blinding helps prevent participants from consciously or unconsciously altering their responses to align with expectations or desired outcomes. By isolating neural responses to the visual stimuli of interest, we mitigate a chief concern about survey evidence: the possibility of biases introduced through task instructions or leading questions.[52]

The third advantage of our approach is the use of RS. It provides a quantifiable measure of the degree of similarity between two stimuli without relying on additional assumptions about how to quantify that similarity.[53] Allowing the brain to provide an analog similarity measure also eliminates the need for researchers to predefine specific criteria or metrics for evaluating similarity, reducing potential biases inherent in subjective evaluations. This technique directly addresses the criticism that trademark surveys force respondents into a simple binary choice as to whether or not two products are confusingly similar when, in reality, the magnitude of any confusion should inform the analysis.

In summary, this methodological framework minimizes biases that can be introduced through task instructions, leading questions, and subjective judgments, allowing us to obtain more reliable and robust findings that may aid legal decision-making in trademark

---

brain region encoding visual object information is interpreted as the neural signature of similarity, without requiring explicit responses from the human participants.

[52] *See, e.g.*, Jack Daniel's Props., Inc. v. VIP Prods. LLC, 143 S. Ct. 1578, 1593 (2023) (Sotomayor, J., concurring) ("Cleverly designed surveys could also prompt such confusion by making consumers think about complex legal questions around permission that would not have arisen organically out in the world.").

[53] Of course the ultimate determination of what measurable amount should be considered probative of visual similarity and, more indirectly, the ultimate issue of consumer confusion has to reside with the trier of fact.

infringement disputes. It is worth noting that these advantages would not be afforded simply by moving a survey into an fMRI scanner. In fact, asking subjects to make similarity judgments within an imaging scanner would introduce the very biases we aim to minimize, potentially leading to the activation of cognitive processes and biases associated with subjective decision-making.

E. Complementing Survey Results with Our Neural Index

Having created a prototype neural similarity index, we set out to develop an experimental test to investigate whether our method could indeed detect biased surveys in trademark litigation. Bias assessment in self-report instruments is notoriously challenging due to the absence of a definitive benchmark. To address this issue, we devised an experimental approach that allows us to manipulate and calibrate bias in a transparent and replicable manner: we created surveys set in a hypothetical legal context, using the same set of reference and competitor products in the fMRI experiment, to collect self-report evaluations of similarity between different products. In these surveys, we incorporated varying degrees of bias, which was intentionally induced to favor either proposed plaintiffs (the maker of REESE'S Peanut Butter Cups and the maker of OXICLEAN laundry detergent) or potential defendants (the maker of "Pieces" and the maker of "OxyClear").

Drawing from documented criticisms of litigation surveys in trademark infringement cases[54] and recent scientific literature on questionable research practices,[55] we employed formats and language commonly seen in trademark cases. We showed that it is possible to reach diametrically opposing conclusions by manipulating elements of the survey including background information, the similarity criteria participants are instructed to use, and the question format. For example, the following variations of the instructions were used in the three versions of the survey, respectively[56]:

- Instruction #1: "For Brand B to be considered a potential copycat of Brand A, it should be the case that Brand B is much more similar than other brands in the same

---

54 Simonson & Kivetz, *supra* note 30.

55 Uri Simonsohn, Leif D. Nelson & Joseph P. Simmons, *P-curve: A Key to the File-drawer*, 143 J. Exp. Psych: Gen. 534 (2014).

56 Admittedly, the biases in trademark surveys submitted as evidence will typically be more subtle than those studied here. We note that testing our neural index on clearly biased survey questions is a necessary first step in this research as it provides a useful positive control. Future work can refine our technique against more subtly flawed surveys.

marketplace as Brand A. A 'yes' judgment for Brand B thus implies that it is considered to be infringing."

- Instruction #2: "To reach a 'yes' judgment, it should be the case that almost all features between the two products should be identical, rather than simply some shared design elements."
- Instruction #3: "Your judgment should be based on the visual elements of the product packages, including but not limited to colors, fonts, overall layout and style, the shape of the package, etc."

Along with other manipulations, we found that the results from version #1 were much more favorable to the plaintiff.[57] In contrast, the results from version #2 were more favorable to the defendant. Finally, the last version yielded results that fell in between the two.

Rather than deferring to survey experts to discuss the relative merits and demerits of each format in circumstances where there is rarely agreement, we examined the relationship between the neural similarity index derived from brain activity in the object-sensitive cortex and the behavioral measures obtained from each survey. In both candies and cleaning products, whereas the neural similarity index showed poor correspondence with the more pro-defendant and pro-plaintiff surveys, it was highly aligned with the more neutral third survey. These findings demonstrate the capacity of this index to capture distinctions between surveys with varying degrees of bias, highlighting its potential as a more objective measure of neural similarity that can clarify data obtained by traditional surveys.

IV. FUTURE APPLICATIONS

A. Neuroimaging and Survey Evidence: Stronger Together

In short, our results demonstrate the possibility of capturing perceived visual similarity—an important component of trademark confusion—by neuroimaging techniques. Despite being a proof of concept, the experiment presents a scenario in which our proposed neural similarity index helps identify biases in survey evidence. With uncertainty as to what defines an acceptable survey, judges have often discounted their evidentiary weight. The neuroscientific measure we developed provides a new opportunity to either buttress or challenge the validity of a survey, particularly of value when each side to a trademark dispute introduces conflicting survey evidence. In this way, neuroscience has the potential to turn surveys into more trustworthy tools in trademark litigation.

---

[57] Zhang et al., *supra* note 8.

This is not to say that RS can tell us exactly when consumers will confuse one stimulus for another. It may be tempting to think of our neural similarity index as the "ground truth" of visual confusion, or at least of consumer perception of visual similarity. However, the mapping between unobservable mental states (such as similarity) and observable brain activities (such as RS driven by visual similarity) remains an assumption that must be continuously tested by ongoing scientific work. As a result, our contribution can be more modestly described as introducing a novel form of evidence for likelihood of confusion. This form of evidence relies on a distinct set of measurements from those that can weaken the validity and reliability of survey evidence. While the neural similarity index can still be susceptible to its own bias and errors, they are unlikely to be the same as those for survey evidence.

Therefore, we emphasize that we are not advocating for the wholesale replacement of surveys with fMRI data in trademark infringement lawsuits. Instead, these two types of evidence are stronger together than apart. Because of the fundamental differences between the processes that produce them, their deviations from the ground truth (if any) will likely be in different directions.[58] For this reason, alignment between two distinct forms of evidence—as seen in the consistency between the neural similarity index and the third neutral survey instruction in our proof-of-concept study—greatly boosts confidence in their validity.

Another potential application of our approach is to serve as a validation method to help improve, and perhaps help promote, best practices for surveys and other empirical methods of measuring trademark confusion.[59] Although methods such as *Ever-Ready* and *Squirt* have been commonly accepted, they provide only loose guidance on how a survey should be designed for a specific dispute, leaving many methodological elements to be determined. In court, these choices often become the target of vehement criticism by expert witnesses from the opposing party. Because assessing the size of their impact on the survey responses is often difficult, such surveys often end up being admitted into evidence.

In parallel, there is a general lack of empirical scholarly research on the robustness of survey responses to design variations. Notable exceptions that examine different variations of survey formats do

---

58 This methodology is often referred to as "triangulation," where multiple independent methods are used to corroborate and cross-verify findings. *See generally*, Todd D. Jick, *Mixing Qualitative and Quantitative Methods: Triangulation in Action*, 24.4 Admin. Sci. Q. 602 (1979).

59 Among other benefits, the costs of fMRI studies can be comparable to those of consumer survey studies. As a reference, our fMRI results, based on sample size typical for visual processing experiments, cost approximately $20,000 for data acquisition, based on 1.5 hours of scanning time per participant at a rate of $650/hr, with an additional $50 per participant for subject renumeration.

show large influences on the survey results,[60] but without a validation method they are less capable of pinpointing the design(s) that generate the most objective responses. Future systematic research may apply our approach to improve and refine the standards for trademark confusion surveys.

B. Other Types of Similarity

Although our proof of concept primarily focuses on visual similarity, the overall approach can be applied to similarity in other sensory domains deemed important in trademark disputes. fMRI and other neuroimaging techniques have been widely applied in the study of neural representations for hearing,[61] touch,[62] taste,[63] smell,[64] and more abstract processes such as memory,[65] language,[66] and social judgment.[67] In parallel, the underlying principle of RS remains consistent across different sensory domains[68] and this generalizability has been empirically documented.[69] It is thus feasible to use an approach like ours to examine similarity between stimuli beyond visual ones. Other methods besides RS for analyzing neural representations from imaging data, e.g., multivoxel pattern

---

60 DeRosia, *supra* note 16, at 620-21.

61 *See, e.g.*, Melissa Saenz & Dave R.M. Langers, *Tonotopic Mapping of Human Auditory Cortex*, 307 Hearing Rsch. 42 (2014) (discussing procedures and analysis for mapping human auditory cortex with fMRI).

62 *See, e.g.*, Burkhard Pleger & Arno Villringer, *The Human Somatosensory System: From Perception to Decision Making*, 103 Progress Neurobiology 76 (2013) (reviewing research on touch perception imaging).

63 *See, e.g.*, Junichi Chikazoe, Daniel H. Lee, Nikolaus Kriegeskorte & Adam K. Anderson, *Distinct Representations of Basic Taste Qualities in Human Gustatory Cortex*, 10 Nature Commc'ns 1048 (2019) (using fMRI to map taste qualities in human brain).

64 *See, e.g.*, A. Fournel, C. Ferdenzi, C. Sezille, C. Rouby & M. Bensafi, *Multidimensional Representation of Odors in the Human Olfactory Cortex*, 37 Human Brain Mapping 2161 (2016) (mapping odor representation in brain with fMRI).

65 *See, e.g.*, Brian D. Gonsalves, Itamar Kahn, Tim Curran, Kenneth A. Norman & Anthony D. Wagner, *Memory Strength and Repetition Suppression: Multimodal Imaging of Medial Temporal Cortical Contributions to Recognition*, 47 Neuron 751 (2005) (applying fMRI and other imaging to examine memory strength in human brain).

66 *See, e.g.*, Kirsten Weber, Morten H. Christiansen, Karl Magnus Petersson, Peter Indefrey & Peter Hagoort, *fMRI Syntactic and Lexical Repetition Effects Reveal the Initial Stages of Learning a New Language*, 36 J. Neuroscience 6872 (2016) (using fMRI to analyze language structures in brain).

67 *See, e.g.*, Adrianna C. Jenkins, C. Neil Macrae, & Jason P. Mitchell, *Repetition Suppression of Ventromedial Prefrontal Activity During Judgments of Self and Others*, 105 Proc. Nat'l Acad. Scis. 4507, 4510 (2008) (examining self-reflection and judgment of others using fMRI studies).

68 *See, e.g.*, Kalanit Grill-Spector, Richard Henson & Alex Martin, *Repetition and the Brain: Neural Models of Stimulus-Specific Effects*, 10 Trends Cognitive Scis. 14 (2006) (explaining that the repetition suppression dynamic occurs across brain regions and under a large range of experimental conditions).

69 Barron et al., *supra* note 45, at 2.

analysis,[70] provide additional methodological flexibility for these needs.

Notably, phonetic similarity—i.e., similarity in how two marks sound—has played an important role in the analysis of likelihood of confusion.[71] The general design of our experiment can be easily adapted to enable the measurement of phonetic similarity between pairs of stimuli, either independently or along with visual similarity.[72] An open question is whether the brand names should be presented visually or auditorily (or both), a choice that merits further study and, of course, may depend on the nature of the specific dispute.[73]

Neuroscientific methods may also inform trademark law by providing a more unified view on what other, potentially more abstract, types of similarity the court should consider. For example, recent research in consumer behavior has demonstrated the effect of theme or conceptual similarity on consumer evaluation of copycat brands.[74] If such theme similarity evokes stable and consistent neural signatures in the brains of representative consumers in the same way that basic featural or auditory similarity does, these data may constitute a powerful argument for more serious consideration of formal protection.

---

70 Multivoxel pattern analysis (MVPA) is a method used in neuroimaging research to understand how information is represented and processed in the brain. It can be used to analyze patterns of brain activity across multiple brain regions, as well as individual brain areas. By examining the unique patterns of activity across voxels (small imaging units in the brain), MVPA can identify specific patterns associated with different mental states or tasks, allowing researchers to decode what someone is seeing, thinking, or experiencing based on their brain activity patterns. *See generally* Tyler Davis & Russell A. Poldrack, *Measuring Neural Representations with fMRI: Practices and Pitfalls*, 1296 Annals N.Y. Acad. Scis. 108 (2013).

71 We are not aware of the introduction of trademark survey evidence on the issue of consumers confusing one sound for another. Nevertheless, the trier of fact has been called upon to evaluate sound similarity in trademark infringement disputes. *See* Pocono Rubber Cloth Co. v. J.A. Livingston, Inc., 79 F.2d 446, 448 (3d Cir. 1935) (SUAVELLE and SWAVEL resemble each other in sound); Bell Publ'g Corp. v. Bantam Doubleday Dell Publ'g Group, Inc., 17 U.S.P.Q.2d 1634, 1637 (E.D. Pa. 1990) (DELL and BELL have confusingly similar sounds).

72 Our proof of concept study actually included phonetic similarity between certain stimulus pairs (e.g., REESE'S vs. "Pieces" and OXICLEAN vs. "OxyClear"), although the experiment was not optimized for systematic analysis, thus we focused only on repetition suppression signals in the visual area of the brain.

73 *See, e.g.*, Virgin Enters. Ltd. v. Nawab, 335 F.3d 141, 149 (2d Cir. 2003) (concluding that plaintiff's advertising on the radio meant auditory similarities controlled consumer perception over unknown visual dissimilarities).

74 Femke van Horen & Rik Pieters, *Consumer Evaluation of Copycat Brands: The Effect of Imitation Type*, 29 Int'l J. Rsch. Mktg. 246 *passim* (2012).

C. Beyond Mark Similarity

Of course, mark similarity is not the only factor investigated in determining likelihood of confusion. Other factors may also be amenable to empirical examination and quantification by neuroscientific methods. The common thread is the focus on measurement of consumer perception, defined as a certain mental state, or set of mental states, shared by representative consumers of a given product. A simple framework for assessing the feasibility of developing neuroscientific measures for these factors consists of two key questions: (1) To what extent does a given factor correspond to one or more well-defined psychological states? And (2) to what extent does the psychological state(s) map onto well-characterized brain activities that can be reliably measured by neuroimaging techniques?

Given this framework, mark strength may be another aspect of the confusion analysis that can profit from neuroscientific techniques. The strength of a mark indicates its ability to distinguish the goods or services of a firm from those of others, and, according to trademark doctrine, higher mark strength warrants more legal protection.[75] As readers of this publication know well, if a mark is so weak (generic) that it is used by consumers to describe a whole class of products, it cannot be considered a trademark at all. Empirical analyses have shown that mark strength is given great weight in the multifactor test of confusion.[76]

In practice, measuring mark strength can be a messy business, as the evidentiary weight of survey evidence may again be discounted,[77] and courts often rely on distal proxies for consumer recognition such as market sales volume or advertising expenses.[78] From a psychological standpoint, the degree to which consumers associate a mark with a particular source, and more generally the meaning of a word, can be conceptualized as a memory phenomenon. More specifically, meanings, concepts, and facts belong to a type of memory called semantic memory, i.e., the general world knowledge shared by people from a specific culture or society.[79] A strong mark will possess an immediate and exclusive association with the source

---

75 J. Thomas McCarthy, McCarthy on Trademarks and Unfair Competition § 11.73 (5th ed. 2023) ("All courts agree that 'stronger' marks are given 'stronger' protection . . . .").

76 Beebe, *supra* note 42, at 1633-34.

77 David H.B. Bednall, Phillip Gendall, Janet Hoek, & Stephen Downes, *Color, Champagne, and Trademark Secondary Meaning Surveys: Devilish Detail*, 102 TMR 967, 970 & n.23 (2012) (criticizing secondary meaning survey designs that test for "association" instead of "identification," leading to "ambiguous results").

78 Lisa Larrimore Ouellette, *The Google Shortcut to Trademark Law*, 102 Calif. L. Rev. 351, 362 (2014).

79 Abhilasha A. Kumar, *Semantic Memory: A Review of Methods, Models, and Current Challenges*, 28 Psychonomic Bull. & Rev. 40, 40-41 (2021).

in consumers' semantic memory, while a generic mark will be tied to the product category instead.[80]

How well can neuroimaging measure such associations? As in our measurement of perceived visual similarity, the opportunity lies in capturing these automatic associations without having to ask questions or prompt explicit judgments. Several streams of cognitive neuroscience literature hint at the possibility of doing so. First, the neuroscience of semantic memory has long been an active area of research,[81] and exciting progress on decoding transient representations of semantic information in specific, yet distributed, regions of the brain has recently been made.[82] Relatedly, research on mental imagery (i.e., activation of mental representations from memory in the absence of a corresponding external stimulus) also reveals that such representations can be decoded using imaging data.[83] Finally, recent work in consumer neuroscience has successfully predicted consumers' associations of brands using fMRI data.[84]

Another concept of great significance in trademark law, albeit outside the scope of the likelihood of confusion factors, is mark dilution. Dilution refers to the weakening or diminishing of the distinctiveness or uniqueness of a famous trademark. It occurs when a mark's reputation or distinctiveness is eroded by the use of a similar or identical mark by another party, even if there is no likelihood of confusion between the marks.

---

80 Courts divide their analysis of trademark strength into two dimensions: conceptual strength and commercial strength. Conceptual strength analyzes the potential for a mark to signal source at the time of its first use and is evaluated for placement along trademark law's distinctiveness spectrum. Commercial strength evaluates actual consumer recognition of the mark at the time of registration or infringement litigation. Here, we are focused on our neuroscience's potential for providing evidence of commercial strength. *See also* Mark Bartholomew, *Neuromarks*, 103 Minn. L. Rev. 521, 534-36, 556-57 (2018) (discussing experiments using fMRI imaging to assess brand familiarity).

81 *See, e.g.*, Jeffrey R. Binder & Rutvik H. Desai, *The Neurobiology of Semantic Memory*, 15 Trends Cognitive Scis. 527 (2011) (proposing semantic memory model based on neuroimaging).

82 *See, e.g.*, Sara F. Popham, Alexander G. Huth, Natalia Y. Bilenko, Fatma Deniz, James S. Gao, Anwar O. Nunez-Elizalde & Jack L. Gallant, *Visual and Linguistic Semantic Representations Are Aligned at the Border of Human Visual Cortex*, 24 Nature Neuroscience 1628 (2021) (reporting separate but adjacent semantic networks for visual versus auditory information); Alexander G. Huth, Shinji Nishimoto, An T. Vu & Jack L. Gallant, *A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories Across the Human Brain*, 76 Neuron 1210 (2012) (identifying a continuous sematic space in the brain representing thousands of categories of objects and actions).

83 Joel Pearson, Thomas Naselaris, Emily A. Holmes & Stephen M. Kosslyn, *Mental Imagery: Functional Mechanisms and Clinical Applications*, 19 Trends Cognitive Scis. 590 (2015).

84 Yu-Ping Chen, Leif D. Nelson & Ming Hsu, *From "Where" to "What": Distributed Representations of Brand Associations in the Human Brain*, 52 J. Mktg. Rsch. 453 (2015).

First, on the question of whether a mark is well known enough to warrant protection against dilution, any neuroscientific evidence of mark strength might also be relevant to assessing whether a particular mark is "famous." Federal dilution law specifically requests consideration of "the extent of actual recognition of the mark."[85]

Second, dilution requires a determination that the defendant's use produces an association in consumers between the defendant's mark and the famous mark that would impair the latter's distinctiveness or harm its reputation. Attempts to use surveys to measure dilution have been controversial, likely because of the difficulties in designing a survey to measure spontaneous associations.[86] Given that the federal dilution statute expressly calls for the trier of fact to assess (among other factors) "the degree of similarity between the mark or tradename and the famous mark,"[87] our neural suppression index could have an immediate impact in dilution matters. Mark dilution also fits well within the conceptual framework of semantic memory, as it is essentially concerned with whether and how the content and strength of the associations of a mark are affected by new associations created by a different mark and/or its marketing actions. While there are likely technical challenges to overcome, especially regarding the complexity of real-world brand associations and the fast, spontaneous nature of mental associations, rapid advances in the cognitive neuroscience of semantic representations indicate the goal may be within the reach of the current generation of cognitive neuroscientists.[88]

V. CONCLUSION

Although other legal subject areas have garnered more attention, trademark law may be uniquely suited to profit from neuroscience. Its tests for validity and infringement hinge on consumer sentiment, and neuroscience can offer probative information on the aggregate perceptions of consumers. Our experiment suggests that, by tracking the phenomenon of repetition suppression in the brain, a neural record of visual similarity can contribute to a more robust portrait of consumer confusion and offer a check on survey results for potential bias. Of course, the mere fact that neuroscience offers new tools for measuring human perception

---

85 15 U.S.C. § 1125(c)(2)(A)(iii).

86 *See* Barton Beebe, Roy Germano, Christopher Jon Sprigman, & Joel H. Steckel, *The Science of Proving Trademark Dilution*, 109 TMR 955 *passim* (2019); Shari Seidman Diamond, *Surveys in Dilution Cases II*, *in* Trademark and Deceptive Advertising Surveys 156, 157-62 (Shari Seidman Diamond & Jerre B. Swann eds., 1st ed. 2012).

87 15 U.S.C. § 1125(c)(2)(B)(i).

88 Sandra M. Virtue & Darren S. Cahr, *Trademarks and the Brain: Neuroscience and the Processing of Non-Literal Language*, 112 TMR 695, 704-05 (2022).

does not tell us exactly how the law should account for those measurements. Confusion is ultimately a legal standard, not a scientific one, and it will be the job of judges, legislators, and advocates to determine how advances in our understanding of the brain may both support and shape trademark law in the future.

RECONSTRUCTING THE TRADEMARK REGISTRY OF MANDATE PALESTINE AND WHAT HISTORICAL DATA CAN REVEAL

By Michael Birnhack[*]

TABLE OF CONTENTS

[*] Professor of Law, Faculty of Law, Tel Aviv University. The reconstructed registry is the product of a team of research assistants, led in two phases by Doron Pe'er and Raz Ashkenazi. I am indebted to their meticulous work, and to the team: Yoav Banai, Adi Ben-Eli, Orly Ben-Moshe, Nita Ben-Oliel, Yonatan Ben-Yosef, Hila Davidi, Yael Iosilevich, Asaf Kramer, Yafit Mamistvalov, Shereen Ounallah, Nimrod Prinz, and Dona Saabni. Bar Ifrah and Eran Toch developed a method to extract the trademarks from PDF files. Oren Ben-Zvi and Michal Lahav of the Brender-Moss Library for Social Sciences, Management and Education at Tel Aviv University developed the search engine, and Pua Eden Holtsman designed the interface. Thanks to Amir Khoury, Roy Kreitner, and Assaf Likhovski for comments on earlier drafts. I acknowledge the support of the David Berg Foundation Institute for Law and History and Israel Science Foundation (ISF) Grant 537/21.

I. INTRODUCTION

Trademark data offers a rich yet underexplored resource for legal historians and business and economy historians, as well as cultural historians and design historians. I use "trademark data" to refer to data about the application for and registration of trademarks. The data are about who applied, when, where, for which kinds (classes) of marks, which applications were withdrawn, accepted, or denied and when, and any other data about the process. Thus, trademark data refer not to the contents of the trademarks but to their meta layer, to data about the contents.

For legal historians, trademark data may reveal yet unnoticed trends beyond the cases that reached litigation. For business historians, the data shed light on corporate plans and strategies. For economic historians, trademark data offer an indicator of innovation and competition and a novel lens through which to evaluate fluctuations in the overall economy or in specific industries. Cultural historians can trace subtle cultural changes, such as consumption trends, and design historians can combine the metadata with the marks themselves, such as for tracing foreign influences or conducting a systematic semiotic analysis rather than offering only anecdotal examples of specific trademarks.

This article makes the case for the importance of utilizing historical trademark data for various fields of study. The main hurdle is that, in most cases, the data are unavailable and need to be gathered and analyzed. This task incurs various challenges and requires some funding and attention. Evaluating trademark data carries some important limitations, and those engaging with these data should be cautious in interpreting them: A key limitation is that not all economic activity concluded in registered trademarks. Thus, the data may reflect some but not all commercial trends. And yet, the potential is worth the effort. This resource can triangulate other traditional resources, such as archival research and legal study of legislation and case law.

To illustrate the potential (and pitfalls) of historical trademark data, I discuss the reconstruction of the trademark registry of Mandate Palestine, from the British registry's debut in 1922 until the Mandate's end and the State of Israel's establishment in 1948.[1] The original registry was lost, leaving this author with little choice

---

[1] In 1917–18, the British conquered the region comprising today's Israel (excluding the Golan Heights), the West Bank, and the Gaza Strip. Initially, they established a military regime, which was replaced by an administrative one in 1920. In 1922, the League of Nations accorded the King with a "Class A" Mandate over Palestine. The Mandate echoed the 1917 Balfour Declaration, including the British statement, "to view with favour the establishment in Palestine of a national home for the Jewish people." The Mandate lasted until May 1948, when the British left, and the State of Israel was established. I use the term "Mandate Palestine" to refer to the British government of the region in 1917–1948.

but to reconstruct it.[2] I discuss the methodological challenges, point to the benefits and opportunities of using the data, and note their shortcomings.

Part II places the research of historical trademark data within a broader research framework. Part III presents the reconstructed trademark registry of Mandate Palestine. I discuss the methodology applied to this task and the challenges encountered during the registry's reconstruction and offer some lessons for similar projects in other jurisdictions. Part IV presents some findings emerging from the reconstructed registry, illustrating the potential of historical trademark data and its limitations. Part V offers some concluding remarks.

II. THE EMPIRICAL TURN

Historical trademark data lie at the intersection of several research threads and topics: (1) A growing interest in trademark data by economics and business scholars conducting contemporary research and historical research in these fields; (2) the (re)emergence of empirical legal studies; (3) first buds of contemporary (rather than historical) legal studies that utilize trademark data; and (4) a rise in interest in the legal history of trademark law. This Part sets the stage with a concise literature review of these building blocks.

A. Trademark Data

The first to conduct data-based empirical studies (not necessarily historical research) in the fields of intellectual property ("IP") were economists. They commenced with patents. Patent registration data present a rich resource to explore innovation trends and detect the lineage of specific inventions, as patents cite prior art.[3] Not all innovation results in patent applications and registrations; hence, economists explored the revealed tip of the iceberg rather than the entire innovation activity in a specific jurisdiction.

Copyright law is more difficult in this regard, as in most countries, the law does not require registration for the legal protection of works of authorship (and, in fact, the Berne Convention, as amended in 1908, prohibits setting formalities as a

---

[2] *Mandate Palestine's Reconstructed Trademark Registry (1917–1948)* (Tel Aviv University, Michael Birnhack, ed., 2023), available at https://en-law.tau.ac.il/MandatePalestineIP.

[3] The pioneering work is Adam B. Jaffe, Manuel Trajtenberg & Rebecca Henderson, *Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations*, 108(3) Q.J. Econ. 577 (1993). For a comprehensive discussion, see Adam B. Jaffe & Manuel Trajtenberg, Patents, Citations, and Innovations: A Window on the Knowledge Economy (2005).

prerequisite for protection).[4] The most notable jurisdiction to maintain such registration is the United States.[5] However, U.S. copyright registration has only limited legal power and, hence, represents only an unknown portion of works of authorship.[6] Only recently, scholars have begun exploring copyright registration data.[7]

Turning to trademarks, as with patent data, economists were the first to explore trademark data. They explore issues such as whether firms' trademark activity is an indicator of innovation[8] or competition,[9] whether there is a correlation between trademark activity and firms' size,[10] as well as examining specific sectors[11] and various other economic issues.[12] Management research has also shown a growing interest in trademark data.[13]

Business historians were next to show interest in trademark data.[14] For example, an early study explored trademarks for export in the cotton industry.[15] Paul Duguid, Teresa da Silva Lopes, and

---

4 *See* Berne Convention for the Protection of Literary and Artistic Works, September 9, 1886, as revised at Paris on July 24, 1971, Art. 5(2) 1161 U.N.T.S. 3.

5 The registry is managed by the Library of Congress and is available at https://www.copyright.gov/public-records/.

6 U.S. copyright law requires registration of works as a prerequisite for undertaking civil action for infringement of U.S. works. *See* 17 U.S.C. § 411.

7 *See* Dotan Oliar, Nathaniel Pattison & K. Ross Powell, *Copyright Registrations: Who, What, When, Where, and Why*, 92 Tex. L. Rev. 2211 (2014); Robert Brauneis & Dotan Oliar, *An Empirical Study of the Race, Ethnicity, Gender, and Age of Copyright Registrants*, 86 Geo. W. L. Rev. 46 (2018).

8 *See*, *e.g.*, Meindert J. Flikkema, Ard-Pieter de Man, Matthijs Wolters, *New Trademark Registration as an Indicator of Innovation: Results of an Explorative Study of Benelux Trademark Data* (Research Memorandum, 2010).

9 *See*, *e.g.*, Po-Hsuan Hsu, Kai Li, Xing Liu & Hong Wu, *Consolidating Product Lines via Mergers and Acquisitions: Evidence From the USPTO Trademark Data*, 57(8) J. Fin. & Quantitative Analysis 2968 (2022).

10 For a literature review, see Mirësi Çela, *The Importance of Trademarks and a Review of Empirical Studies*, 4 Eur. J. Sustainable Development 125 (2015).

11 *See*, *e.g.*, Jasper Grashuis, *Branding by U.S. Farmer Cooperatives: An Empirical Study of Trademark Ownership*, 5 J. Coop. Org. & Mgmt. 57 (2017).

12 *See*, *e.g.*, Eric J. Iversen & Sverre J. Herstad, *Dynamics of Regional Diversification: A New Approach Using Trademark Data*, 56(2) Reg'l Studies 276 (2022). For a literature review, see Shukhrat Nasirov, *The Use of Trademarks in Empirical Research: Towards an Integrated Framework* (Working paper Nov. 20, 2018).

13 For a literature review, see Carolina Castaldi, *All the Great Things You Can Do with Trademark Data: Taking Stock and Looking Ahead*, 18(3) Strategic Org. 472 (2020).

14 *See* Trademarks, Brands, and Competitiveness (Teresa Da Silva Lopes & Paul Duguid, eds., 2010); Montserrat Llonch-Casanovas, *Trademarks, Product Differentiation and Competitiveness in the Catalan Knitwear Districts during the Twentieth Century*, 54(2) Bus. History 179 (2012); Sáiz Patricio & Fernández Pérez Paloma, *Catalonian Trademarks and the Development of Marketing Knowledge in Spain, 1850–1946*, 86 Bus. History Rev. 239 (2012).

15 D. M. Higgins & Geoffrey Tweedale, *The Trade Marks Question and the Lancashire Cotton Textile Industry, 1870–1914*, 27(2) Textile History 207 (1996).

John Mercer examined trademark registrations in France, the United Kingdom, and the United States, asking whether trademarks are a proxy for innovation in marketing.[16] Trademark data assisted others in analyzing the business strategies of British multinational companies.[17] For example, World Bank economists examined the worldwide distribution of trademarks, finding an asymmetry of ownership, with most trademarks registered by firms in industrialized countries.[18]

These studies apply economic and business lenses and offer important insights about how to approach trademark data, their potential to shed new light on economic processes, and some caveats. They treat the law as fact and do not typically inquire about changes in the law itself, its underlying rationales, doctrines, or effects.

B. Empirical Legal Studies

Moving from economic and business studies to the law, we notice the reemergence of empirical legal studies. More than a century ago, American legal realism introduced social sciences to the study of law, pushing aside doctrinal analysis. During the twentieth century, the realist shift evolved, splitting into several branches, such as critical legal studies and law and economics, and later additional approaches, such as law and literature. Many of these approaches utilize data in various ways, such as relying on available statistics related to the topic of inquiry. Law and society scholars have used interviews, ethnographic tools, and other qualitative empirical methods for some time.

Empirical legal studies first appeared alongside legal realism[19] and have reemerged in earnest in the past two decades. We now witness the consolidation of a field of empirical legal studies.[20] The unifying feature is the interest in the methodology that may apply to diverse legal topics.

No research approach is void of underlying biases. The challenge is to recognize them early on. In some cases, this is a proclaimed

---

16 Paul Duguid, Teresa da Silva Lopes & John Mercer, *Reading Registrations: An Overview of 100 Years of Trademark Registrations in France, the United Kingdom, and the United States*, *in* Trademarks, Brands, and Competitiveness 9 (Teresa Da Silva Lopes & Paul Duguid, eds., 2010).

17 Teresa da Silva Lopes & Mark Casson, *Brand Protection and the Globalization of British Business*, 86 Bus. History Rev. 287 (2012).

18 Eugenia Baroncelli, Carsten Fink & Beata Smarzynska Javorcik, *The Global Distribution of Trademarks: Some Stylised Facts* (World Bank, Policy Research Working Paper No. 3270, 2004).

19 *See* John Henry Schlegel, American Legal Realism and Empirical Social Science (1995); Michael Heise, *The Past, Present, and Future of Empirical Legal Scholarship: Judicial Decision Making and the New Empiricism*, 2002 U. Ill. L. Rev. 819 (2002).

20 *The Journal of Empirical Studies* first appeared in 2004. A first conference took place in 2006, and today, numerous such conferences take place.

motivation, such as in feminist studies, where scholars build on various premises about the law's masculinity, theories of equality, and social practices. Empirical methodologies are more subtle about their ideological motivations and may, at first sight, seem neutral. However, such underlying assumptions are always at play, and the prudent researcher should strive to be aware of them. For example, not all things are easily quantifiable; hence, decisions regarding what to examine and how to do so are as crucial as determining what is left outside. The researcher inevitably makes numerous decisions about which data to collect and which to omit, which may reflect hidden biases. This much-needed caution is not unique to empirical legal studies. Anticipating the legal historical approach, contemporary historians, too, are keenly aware of the silence of the archives; namely, whereas archives may contain much documentation of the past, much more may be missing.[21] This lesson is applicable to the use of trademark data for historical research.

The rise of empirical legal studies, along with growing methodological capabilities and new digital and computational tools, pave the way for lawyers to engage with trademark data.

C. Trademark Data in Law

Following the economists' interest in trademark data and working within the empirical turn, IP law scholars have also noticed trademark data. Jeremy Sheff constructed datasets of trademarks in Canada and Japan, enabling descriptive analyses of the trademark systems.[22] Amir Khoury compiled and analyzed trademark data about registration in Arab countries in the latter part of the twentieth century, before and after the entry into force of the Agreement on Trade-Related Aspects of Intellectual Property Rights (the "TRIPS Agreement"), focusing on whether applicants were residents or non-residents and examining gaps.[23] Ilanah Fhima and Catrina Denvir empirically analyzed a specific trademark doctrine (*Likelihood of Confusion*).[24] Deborah Gerhardt and Jon Lee explored U.S. trademark data from 1981[25] and summarized prior trademark empirical studies.[26] Barton Beebe and

---

21 *See* David Thomas, Simon Fowler & Valerie Johnson, The Silence of the Archive (2017).

22 Jeremy N. Sheff, *The Canada Trademarks Dataset*, 18 J. Empirical Legal Studies 908 (2021); Jeremy N. Sheff, *The Japan Trademarks Dataset: A First Analysis* (Foundation for Intellectual Property, Institute of Intellectual Property, Japan, 2017).

23 Amir H. Khoury, *Measuring the Immeasurable: The Effects of Trademark Regimes: A Case Study of Arab Countries* 26 J. L. & Commerce 11 (2006).

24 Ilanah Fhima & Catrina Denvir, *An Empirical Analysis of the Likelihood of Confusion Factors in European Trade Mark Law*, 46 IIC – Int'l Rev. Intell. Prop. & Competition L. 310 (2015).

25 Deborah R. Gerhardt & Jon J. Lee, *A Tale of Four Decades: Lessons from USPTO Trademark Prosecution Data*, 112 TMR 866 (2022).

26 *Id*. at 875-78.

Jeanne Fromer studied EU trademark registrations to decipher globalization processes and market integration.[27]

These studies offer important lessons about the practice of the law and enable us to observe legal gaps and the law's limitations. Some of these studies examined not only contemporary trademark practices, extending their analysis to older practices such as Gerhardt and Lee's study, but these are not necessarily historical studies.

D. Trademark Legal History

In the meantime, the legal history of various IP fields has drawn growing interest. Thus far, copyright law has attracted most of the attention, with patent law second and trademark history lagging, with only scant scholarly attention. The few available works focus on the law in Western economies.

The first wave of legal histories of trademark law pointed to ancient practices of using marks, such as marking cattle and pottery,[28] going as far back as the biblical story of Cain,[29] then progressing from ancient times through the Middle Ages to the law at the time of writing. Medieval practices indicate some use of marks for goods but not trademarks in the modern sense. These marks often operated within the guild system, and in many cases, the marks were required rather than initiated by the manufacturers and traders. Marking was a regulatory mechanism. Indeed, to this day, we find some industries where manufacturers are required to mark their goods, such as in the pharmaceutical industry, to achieve various public interests rather than specific commercial interests.

With the demise of the Middle Ages' guild system, markets opened to competition. The major change in trademark use transpired in the 19th century. Sidney Diamond pointed to several factors that brought about the change: Modern manufacturing methods replaced handwork; production was concentrated in larger units, which required developing distribution methods; and advertising was introduced to acquaint the public with the goods, precipitating trademarks to identify the source of the goods.[30] Thomas Drescher added the "environment of global markets, free

---

27 Barton Beebe & Jeanne C. Fromer, *The Future of Trademarks in a Global Multilingual Economy: Evidence and Lessons from the European Union*, 112 TMR 902 (2022).

28 *See*, *e.g.*, Edward S. Roger, *Some Historical Matter Concerning Trade Marks*, 9 Mich. L. Rev. 29 (1910); Abraham S. Greenberg, *The Ancient Lineage of Trade-Marks*, 33 J. Patent Office Society 876 (1951); Gerald Ruston, *On the Origin of Trademarks*, 45 TMR 127 (1955); Sidney A. Diamond, *The Historical Development of Trademarks*, 65 TMR 265 (1975).

29 Greenberg, *supra* note 28.

30 Diamond, *supra* note 28, at 280-81.

competition, and mechanized production,"[31] especially highlighting advertising, which enabled the creation of a "product identity."[32] These explanations about de-monopolization, industrialization, and (early forms of) globalization gave rise to the modern trademark system and fit the market functions theory of trademark law. These changes explain the timing of the arrival of the first modern trademark laws in the second half of the 19th century, with the advancement of the industrial revolution.

Most historical studies focused on the leading jurisdictions of the time, namely, Great Britain, with its imperial scope and powers,[33] and the United States.[34] In the colonial context, few historical accounts addressed trademark law, with the notable exception of the self-governing dominion of Colonial Australia.[35]

E. Historical Trademark Data

Tying the threads of the empirical turn in legal studies, the growing interest in trademark data, and the rise in IP legal history, the next step is to analyze historical trademark data through a legal lens.[36] So far, we have very few such studies, with Amanda Scardamaglia pioneering this thread. She studied colonial trademarks in Australia by examining a sample of colonial Australian trademarks.[37] As for Mandate Palestine, no such studies are available.[38]

Why has historical trademark data been relatively neglected thus far? I suggest several explanations. *First*, the empirical turn in legal studies is still in the making. Empirical studies require training in statistical methods or funding to acquire the needed assistance and are not obvious to text-savvy scholars. This may be changing today with a new generation of scholars acquiring such

---

31 Thomas D. Drescher, *The Transformation and Evolution of Trademarks: From Signals to Symbols to Myth*, 82 TMR 301, 321 (1992).

32 *Id*. at 322-24.

33 Lionel Bently, *The Making of Modern Trade Mark Law: The Construction of the Legal Concept of Trade Mark (1860–1880)*, *in* Trade Marks and Brands: An Interdisciplinary Critique 3 (Lionel Bently, Jennifer Davis, Jane C. Ginsburg, eds., 2010).

34 Robert G. Bone, *Hunting Goodwill: A History of the Concept of Goodwill in Trademark Law*, 86 B.U. L. Rev. 547 (2006).

35 Amanda Scardamaglia, Colonial Australian Trade Mark Law: Narratives in Lawmaking, People, Power and Place (2015).

36 A librarian noted the value of patent and trademark data for various issues, including the historical value of patent data. *See* Roger V. Skalbeck, *New Research Uses for Patent and Trademark Data*, 27(2) Legal Info. Alert 1, 4 (Feb. 2008).

37 Scardamaglia, *supra* note 35.

38 An LLM thesis offered a general overview of IP law in Mandate Palestine. *See* Ihab G. Samaan, A Historical View of Intellectual Property Rights in the Palestinian Territories (LLM Theses and Essays, paper 49, University of Georgia, 2003), available at http://digitalcommons.law.uga.edu/stu_llm/49.

capabilities and new tools emerging, such as various Artificial Intelligence ("AI") technologies. *Second*, trademark law is more specific and confined than other legal fields, such as criminal law, where we find a thriving adjunct field of criminology. Dealing with trademark data necessitates at least some familiarity with the field and its unique features, especially when dealing with historical data. A design historian, for example, may not be aware of trademark registrations or consider them as potential resources for research. *Third*, whereas patent law has attracted the attention of economists of innovation, the explanatory power of trademark data––namely, their ability to reflect social and economic processes—may be less evident. *Fourth*, in many cases, historical trademark data is not readily available.

It is time for this picture to change. With the rise of empirical legal studies and the increase of researchers trained in empirical methodologies, the computerization of trademark registrations, and the digitization of archives, as well as the increasing availability of AI tools,[39] many of these hurdles are now lower, opening the door to new studies. However, some challenges persist. To illustrate the benefits and risks of studying historical trademark data, I discuss the reconstruction of Mandate Palestine's trademark registry.

III. RECONSTRUCTING MANDATE PALESTINE'S TRADEMARK REGISTRY

Having completed a study on copyright law in Mandate Palestine,[40] I turned to trademark and patent histories. The law in place at the time was the 1921 Trade Marks Ordinance that introduced the British trademark registration system into the region,[41] but the original registry was missing. An inquiry with the Israeli Patent and Trademark Office ("PTO"), the successor to the British PTO, revealed that the original Mandate registry, which survived more than one war and the transition from the British to the Israeli government, was lost. There is some speculation about this loss, but for whatever reason, the registry was missing. The Israeli PTO has some data and offers an open, public search engine containing all applications that were valid under Israeli law:[42] It contains 1618 trademarks submitted during the British period and continued upon Israel's establishment. Some of these have been

---

39 *See* Stephen Petrie et al., *TM-Link: An Internationally Linked Trademark Database*, 53(2) Au. Econ. Rev. 254 (2020).

40 Michael D. Birnhack, Colonial Copyright: Intellectual Property in Mandate Palestine (2012).

41 Trade Marks Ordinance 1921, *Official Gazette* 57 (Dec. 15, 1921).

42 *See* Israel Patents Office – Trademark Search Online, available at https://trademarks.justice.gov.il/TradeMarkSearch/TradeMarkSearch?lang=en.

renewed over time, and today (December 2023), 494 are still valid. However, all other registry entries regarding applications submitted during the British Mandate that had expired before the establishment of Israel disappeared.

In the absence of the original registry and with the partial Israeli resource, it became clear that reconstructing the original registry was essential to studying Mandate Palestine's trademark data. This Part outlines the process and points to some of the challenges encountered, as well as to the mistakes I have made.

A. Reconstructing the Registry

To compensate for the disappearance of the official registry, I sought alternative resources. The obvious one was the British Government's official publication, the *Official Gazette*, later renamed *The Palestine Gazette* (together, "*The Gazette*"). The first task was to gather all trademark applications for the time and place. When I embarked on this project, the *Gazette* publications were available only in print. Luckily, my law school's library had all copies, the vast majority in English, as well as Hebrew translations.[43] A first lesson for those embarking on similar projects is easy to state, though not always easy to follow: Pause and plan ahead.

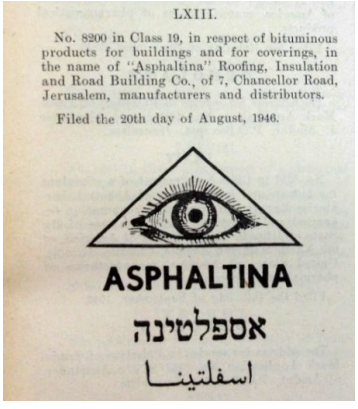
A team of research assistants reviewed 311 *Gazette* publications spanning over 26 years (and later, Israeli official publications were added by a digital search) page by page, as the trademark applications were not published in any consistent manner and appeared sporadically, among other official notices. The process was long, tedious at times, and required some funding. The task would be easier today, as all *Gazettes* have been digitized. The research assistants photocopied each application. The layout of the *Gazette* was not user-friendly, as it appeared in two columns, often with an application beginning at the end of one column and continuing at the top of the next column. A second lesson is straightforward: If the raw resources are not digitized, try to digitize them first in a machine-readable format. I repeatedly had to revert to the original documents.

Once the trademark applications were gathered, we sought to maximize the data extracted from each application. Figure 1 illustrates the extraction.

---

[43] Today, these are available at the Yale Arabic and Middle Eastern Electronic Library, at https://findit.library.yale.edu/catalog/digcoll:2845214, and "NEVO" (a commercial legal database in Hebrew, with subscription, offering various search tools).

Figure 1: Extracting Trademark Data


LXIII.

No. 8200 in Class 19, in respect of bituminous products for buildings and for coverings, in the name of "Asphaltina" Roofing, Insulation and Road Building Co., of 7, Chancellor Road, Jerusalem, manufacturers and distributors.

Filed the 20th day of August, 1946.

ASPHALTINA

אספלטינה

اسفلتينا

From this application, we extracted the following data: application no. (8200), class (19), application date (Aug. 20, 1946), applicant name (Asphaltina Roofing, Insulation and Road Building Co.), place of application (Mandate Palestine), city (Jerusalem), date of publication (May 8, 1947).

We added our own coding: We classified the mark as containing a designed mark and text, we added the languages on the mark (English, Hebrew, Arabic), and coded the graphic contents as "triangle, shape, eye."

Subsequent *Gazette* publications indicated that the application was accepted on Nov. 20, 1947. Here is a third lesson: Extract maximum information. Initially, I downplayed some information, which meant that later we had to return to the raw material and extract additional layers of information. Where possible, classify and code the contents as you go. For example, while we coded the marks' contents, only at a later point did it occur to me to differentiate between word marks, illustrative marks, and combinations; for the word marks, only later did it occur to me to identify the languages used (English, Hebrew, and Arabic being the dominant languages, in this order, with a long tail of other languages). Once again, hopefully, in the not-too-distant future, AI tools will be able to perform many of these tasks.

Extracting data is tedious but worthwhile. The data are objective. The research population comprises all published applications, namely, N = All (but as I explain below, it does not cover all marks used in commerce). Once collected, the published applications are easily processed. I added interpretive layers as we proceeded. For example, we coded the marks' contents. In some cases, it was an easy task: "crown," "camel," "sun," "oranges." Yet, in other cases, deciphering the contents required closer attention: Is it a horizontal crescent or an illustration of a cognac glass?[44] Each mark was reviewed by more than one team member to ensure consistency. Cautious analysis notwithstanding, coding the contents was inevitably subjective, and we may have made mistakes.

---

44 Palestine Trademark (PTM) #1400 (March 1, 1928).

Another subjective assessment was identifying the national identity of the applicants. This may be less interesting for some jurisdictions, but in the intense national atmosphere of the British Mandate, it was an important element to explore. I devised a short checklist to identify the local applicants. In many cases, the name was a strong and sufficient indication of the applicant's Jewish, Muslim, or Christian identity. Noam Levinstein is a Jewish name,[45] and Mohammed Chams El Dine El Dabbagh was characterized as Muslim.[46] Some names are less obvious, and some corporate names are less indicative of their national origin. Accordingly, a second criterion was the applicant's place of residence. An applicant from Bnei Brak, a Jewish city near Tel Aviv, was bound to be Jewish,[47] and Sulphur Quarries Ltd. from Gaza, an almost all-Muslim city, indicated otherwise.[48] Yet, there were mixed cities, especially Jaffa, Jerusalem, and Haifa. The languages used in the trademarks provided a third criterion, as Arab applicants did not use Hebrew. Having applied the previous criteria, the list of "unknowns" was narrowed substantially, enabling a one-by-one search for the remaining applicants' backgrounds. For example, for application No. 8200, featured in Figure 1, the company's name is in English (Asphaltina Roofing, Insulation and Road Building Co.), the city—Jerusalem—was a mixed city, and hence, these criteria are inconclusive. The language mix suggests that the owner was not Muslim, and the use of "Co." rather than the local "Ltd." indicated that this was a local branch of a Boston-based company.

B. Challenges

The registration system posed its own challenges. Here is another lesson: Study the relevant trademark procedure that applied at the place and time under review to the extent possible, as this may shed light on some mysterious issues you are likely to encounter. In the case of Mandate Palestine, I encountered four main challenges. For some of these, it took a while to recognize and more time to sort out: deciphering the application numbering system, missing applications, trademark classes, and the transitions—first, the initial introduction of the British system, and then, the transition into the Israeli system.

As for numbering, the British began with numbering applications per class, irrespective of other classes. Thus, in thirty of the fifty classes, we found application No. 1. This means that organizing the applications according to the assigned numbers

---

45 PTM #1052 (Sept. 29, 1926).

46 PTM #1489 (July 17, 1928).

47 PTM #3763 (June 8, 1935), by Siso Chemical Factory.

48 PTM #4509 (July 2, 1937).

would misrepresent the timeline, such as when asking: Which was the first application? This is yet another reason for maximizing extracted data from each application, as the application dates may be useful in this regard. Two years into this numbering system, in 1924, the British shifted to a unified, consecutive system irrespective of the classes, which better reflects the timeline. However, in 1928, they renumbered the first 592 applications. Thus, for these 592 applications, we have two numbers—the initial one and the reassigned one. For example, the American company Fairbanks, Morse & Co. submitted a trademark application on May 10, 1922. Initially, it was assigned No. 1 in Class 22 and then reassigned to No. 131. The closest trademark under the renumbered system was No. 128 (by a Swedish company, Aktiebolaget Radius), submitted on June 27, 1922, six weeks later. Again, the way to overcome the complexity of such an application was first to recognize it, then search for official corrections, also published sporadically in the *Gazette*. Extracting more data, especially dates, enabled us to sort this issue and avoid mistaken conclusions based on application numbers alone.

A related issue was missing applications. Reviewing the extracted data indicated that we had gaps. For example, we found application No. 1016 (by the British company of Coleman and Company, Ltd.), but the subsequent one was No. 1018 (by the American Standard Oil Company), skipping application No. 1017. This is an indication that an application was submitted and received a number but had not reached publication. Thus, other than the existence of an application, the reconstructed registry is silent: We do not know who the applicants were or what was the sought mark. Based on the last application number submitted during the British Mandate (No. 9778) and comparing it with the data we had regarding 7904 applications and another 45 submitted prior to 1922, I concluded that 1919 applications were discontinued, comprising 19.6 percent of the submitted applications. This is not a negligible share, and any conclusions derived from the dataset should consider this issue. For example, one of the findings emerging from the reconstructed registry was that Jewish-owned trademarks outnumbered Arab Palestinian trademarks. Information about the missing applications may have changed this balance.

For some of the missing applications, their absence was due to the transition periods, as discussed below. For others, however, their absence was because some applicants did not pay the fee, did not submit all documents properly, or later, at the outbreak of World War II, a special emergency Ordinance intervened, instructing the cessation of enemy application processing.[49] To better understand

---

[49] Patents, Designs, Copyright and Trade Marks (Emergency) Ordinance 1939, 973 *Palestine Gazette* 1485.

the missing applications, familiarity with the law was critical, as well as organizing the data systematically to determine whether there was an external reason. In many cases, later official notices indicated changes and corrections that shed some light on these missing applications.

A third challenge was the result of the 1921 Ordinance replaced by the 1938 Ordinance, which came into effect in 1940.[50] The former legislation classified trademarks into 50 classes, with 4601 marks registered under this classification; the latter Ordinance reshuffled the classification into 34 classes with 3258 trademarks. The reclassification required us to separate the data into two clusters; otherwise, someone examining applications submitted under Class 15 for glass (under the 1921 Ordinance classification) without realizing the reclassification may mistakenly include applications from another industry, as Class 15 under the 1938 Ordinance referred to musical instruments. Having realized this issue and to facilitate industry-based analysis over the entire period, we added our own categorization of the industries, creating a common denominator bridging the two periods. The fifty classes under the 1921 law and the thirty-four classes under the 1938 law were grouped into nine categories, such as "professional tools," "food," and "clothing." As this classification is subjective, others may have offered different categories.

Finally, the transition periods posed their own challenges, with the first (from Ottoman to British) posing more of a challenge than the second (from the British to the Israeli system). The Ottoman Empire controlled the entire region for four centuries, ending with the British conquest of the Middle East, including Palestine, in late 1917 and early 1918. The Ottomans had trademark laws in place based on a French law from 1857.[51] Local trademark registration was carried out in Istanbul rather than locally. The Istanbul archives have yet to be explored to determine whether such applications existed. We can assume that since the outbreak of World War I in 1914 and until 1918, hardly any trademark activity transpired in the area, as the war was quite devastating for the entire region. Then, in 1919, the British published an official notice allowing the re-registration of previously registered marks. It is unclear where, when, and how such applications were submitted prior to the entry into force of the 1921 Ordinance and under which law.[52] The reconstructed registry indicates forty-five marks assigned numbers preceded by an "X" with little data on the

---

50 Trade Marks Ordinance, 1938, 843 *Palestine Gazette*, Supp. 1, at 126 (Nov. 21, 1938).

51 Distinctive Marks Act 1871; Regulation on Trademarks concerning Industrial Products and Commercial Goods 1888. For the French connection, see Hasan Kadir Yilmaztekin, *The Legislative Evolution of Copyright in the Late Ottoman Empire*, 17 J. Intell. Prop. L. & Practice 45, 52 (2022).

52 Public Notice 136, Registration of Trademarks, *Official Gazette* (Nov. 16, 1919).

applicants and without the contents of the trademarks. The names indicate mostly British companies, such as The Gramophone Company, John Yates & Co., William Gossage & Sons, and the British-American Tobacco Co. This was a strong indication that these marks were submitted during the British administration prior to the entry into force of the 1921 Ordinance and likely based on the 1919 Notice. Data from twenty years later enabled us to fill some gaps, as some of these applications were renewed, revealing the original date of application between 1919 and 1921.

A related issue concerned the mystery of thirty-three applications for which we had little data: The *Gazette* did not cite application numbers or include the contents of the marks. The dates indicate that these were the first batch of trademark applications submitted after the entry into force of the new law. It seems that these applications were discontinued. Learning the workings of the British registration practices of the time and based on some later official notices, the most plausible explanation I can offer is that at the beginning, the British were unsure how to handle applications that did not meet all requirements, such as unpaid fees or missing documents, so they listed them in the registry but did not assign them numbers. Later, such cases—the noted missing applications—indicate a different practice: An application received a number, but the application was not published until all formalities were met.

The British-Israeli transition was easier to decipher. We found 596 applications submitted during the Mandate that the new Israeli PTO continued. The data were extracted from the digitized Israeli Official *Gazettes* of 1948–1951 in Hebrew. The new registrar retained the British numbering system and continued it rather than beginning anew; it retained the classification under the 1938 Ordinance and examined pending applications. Thus, from a procedural, technical perspective of trademark practices, there was little change.

Once the registry was reconstructed and the challenges were identified and solved, university librarians developed a search engine adapted from existing catalogs to integrate the reconstructed registry with the library's overall search options. To extract the images from the scanned *Gazette* publications, engineers adapted other tools.[53] Numerous trials and corrections ensued. The registry is now in the air, open to all to use under a Creative Commons license.

---

[53] *See* Bar Ifrah, Michael Birnhack & Eran Toch, *Trademarks Extraction and Classification from the British Mandate's Palestine Gazette* (2023), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4498335.

C. Intermediate Recommendations for the Reconstruction of Trademark Registries

The above carries some recommendations for those interested in engaging in similar reconstruction projects elsewhere. *First*, study the procedural aspects of the law, as well as its substance. For example, the initial duration of protection is handy, as subsequent events, namely renewal of a registration, may shed light on the initial application. *Second*, plan ahead to the extent possible. Search for available resources, digitize whatever you can, and search for AI tools that could assist you. *Third*, extract everything you can—contents and metadata. *Fourth*, search for abnormalities in the data, such as missing applications, changes in the numbering system, and the like. The earlier you resolve such issues, the easier it will be to use the data later. The abnormalities may carry their own story, worthy of examination. *Fifth*, code the data. Apply intercoder reliability tests, namely, conduct pilots to ensure that your team members code data similarly and devise a review mechanism to minimize discrepancies and mistakes. Finally, accuracy and consistency are always valuable to maintain. For example, reckoning in advance the best pattern to denote dates (e.g., DD/MM/YYYY or MM/DD/YYYY) will save much time when you begin using the spreadsheet.

IV. SOME FINDINGS, UPSIDES AND DOWNSIDES

Once reconstructed, the registry is ready to explore. This Part presents some of the initial findings from the reconstructed trademarks registry of Mandate Palestine. I have elaborated on these issues elsewhere,[54] and accordingly, the purpose here is to highlight the methodological aspects of such endeavors. I begin with some findings and their benefits and then point to some shortcomings of such data and related caveats.

A. Findings

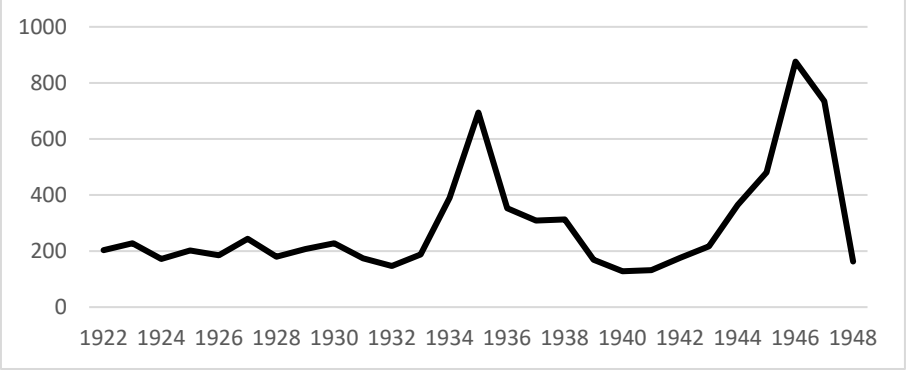
The reconstructed registry contains metadata, enabling us to identify patterns along the parameters used in coding and processing the raw data. The immediate parameters are the time of application, the applicant's identity (including their country of origin), the trademark class, and various combinations of these parameters. Additional metadata may relate to the duration of reviewing the applications and the representation by agents. Alongside the data concerning the applications, we have the trademarks themselves, ready for individual inspection and

---

[54] Michael Birnhack, *Colonial Trademark: Law and Nationality in Mandate Palestine, 1922–1948*, 46(1) Law & Social Inquiry 192 (2021).

systematic semiotic analysis. I discuss the data-based processing options. Here are some of the overall findings revealed by the reconstructed registry, which otherwise would be difficult to observe.

Figure 2: Overall Trademark Registration in Mandate Palestine (1922–1948)



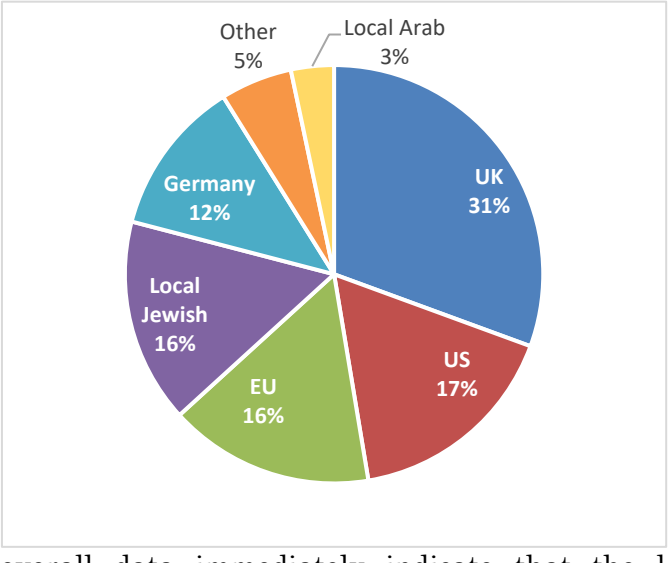
The overall picture of trademark registrations during the Mandate shows the value and relevance of trademark data. Historians of the Mandate can easily recognize the timeline: a sluggish economy during the 1920s; the fifth wave of Jewish immigration from Germany in the early 1930s following the rise of the Nazis to power (with many of the immigrants being traders, closely familiar with trademarks); the Arab Revolt of 1936–39 with a substantial economic slowdown; the beginning of World War II in late 1939; and the economic boom of Mandate Palestine during the war[55] (the figures for 1948 reflect registrations only until May 14, when the Mandate ended). The trademarks offer a strong indication of the economic situation.

The reconstructed registry enables us to break down the data according to the applicants' country of origin: Figure 3 distinguishes between Jewish and non-Jewish applicants from Mandate Palestine and Europe; for convenience, I separated Germany, which had a substantial share.

---

55 For Mandate Palestine's economy, *see e.g.*, Nachum T. Gross & Jacob Metzer, *Palestine in World War II: Some Economic Aspects*, *in* The Sinews of War: Essays on the Economic History of World War II 73 (Geofrey T. Mills & Hugh Rockoff, eds., 1993); Jacob Metzer, *The Economy of Mandatory Palestine: Reviewing the Development of the Research in the Field*, *in* Economy and Society in Mandatory Palestine 1918–1948, 7 (Avi Bareli & Nahum Karlinsky, eds., 2003) (Hebrew).

Figure 3: Trademark Applicants in Mandate Palestine, 1922–1948



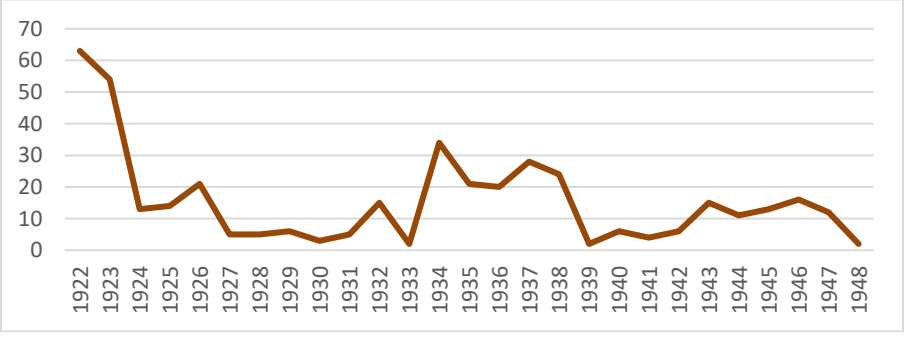
The overall data immediately indicate that the legal tool primarily served foreigners rather than local traders and that within the local market, Jewish traders used the system far more than the Arab traders. External sources could add another layer, such as adding data regarding the composition of the local population; this would show that the Jewish-Arab trademark gap was even greater than the trademark data indicate. This finding supports the divided economy thesis of Mandate Palestine, an issue heatedly debated among the Mandate's economic historians.[56] However, as I explain in the next section, trademarks were not relevant for all industries. Given the voluntary nature of trademark registration, we should be cautious not to conclude that the Arab population did not use marks; the accurate conclusion would be that Arab traders used the British trademark system less than the Jewish traders.

A final example is from a particular industry. Figure 4 presents the timeline of the registered trademarks in the tobacco industry.

---

[56] *See, e.g.*, Barbara J. Smith, The Roots of Separation in Palestine: British Economic Policy 1920–1929 (1993); Jacob Metzer, The Divided Economy of Mandatory Palestine (1998). For a critique, see Zachary Lockman, Comrades and Enemies: Arab and Jewish Workers in Palestine, 1906–1948 (1996).

Figure 4: Tobacco Trademarks During Mandate Palestine 1922–1948



The data indicate that tobacco deviated from the overall picture of trademark applications during the Mandate (see Fig. 2). We see a strong beginning in the early 1920s, then a slowdown to a halt until the mid-1930s, a small recovery and then again, sluggish activity. Adding additional data layers, such as the applicants' identity, sheds more light on this picture.

Extensive trademark activity typically indicates the level of competition in each market. This is an important benefit of resorting to such data. Indeed, a high level of activity may signal robust competition, whereas the opposite—little trademark activity—may indicate a lack of competition in the local market. This was the case with the tobacco industry. In 1883, the Ottoman Government granted the Régie Company a monopoly over the tobacco business in Palestine, Trans-Jordan, and Iraq. The British canceled this concession in 1921.[57] The opening of the market for competition coincided with the coming into effect of the 1921 Trade Marks Ordinance. Indeed, we see numerous applications submitted shortly thereafter (Fig. 4). In the following years, we see a substantial decrease in applications. Additional research into the industry, based on other sources, indicates two related reasons for this decrease. First, the British began regulating the tobacco market in all its segments, from growing and packing tobacco to selling tobacco products.[58] Second, a series of mergers and acquisitions resulted in foreign companies assuming a substantial share of the local market.[59] The trademark data reflect these changes quite clearly.

---

57 The Régie Company demanded compensation for the cancellation of their monopoly, which resulted in long negotiations with the British government. *See* Israel State Archive, M/80/65.

58 See Tobacco Ordinance, 1925.

59 *See, e.g.*, Deborah S. Bernstein, Constructing Boundaries: Jewish and Arab Workers in Mandatory Palestine 125 (2000).

B. Shortcomings

Exploring the data should be handled cautiously, as with all datasets under investigation. I point to five main shortcomings and related caveats: (1) the limitations of the examined dataset, (2) the extent to which trademarks capture commercial activity, (3) the modernized-economy bias of trademarks, (4) a registration bias, and (5) an interpretation bias.

First, for researchers, learning the specific characteristics of the scrutinized dataset is crucial. In Part III, I discussed the challenges I encountered in reconstructing the registry. In the case of Mandate Palestine, the challenges instruct us to be cautious in discerning conclusions about the timeline from the application numbers before 1928 due to the renumbering and reassignment of discontinued applications.

Second, trademark data reflect some commercial activity but not all. Many sales were transacted under the trademark radar. For example, trademarks are less important in open markets—e.g., buying vegetables in a weekend farmers' market. When a customer can closely examine the vegetables, their origin matters less. In such open markets, customers and growers-sellers can interact directly; questions can be asked and answered, thus satisfying the informational roles that trademarks typically have. A century ago, Frank Schechter famously described the trademark's function as the manufacturers' long arm "reaching their hands over the retail tradesman's shoulder, and offering their goods in their own name to the customer."[60] Rephrased in these terms, in open markets and similar situations, such as peddlers, there was no need for stretching the trademark's arm.

A related third shortcoming is that trademark data can better capture modernized economies and inevitably overlook traditional modes of production. When manufacturing is local, the customers are familiar with the origin of the products firsthand. They have all the information they need about the product's origin: They know who the manufacturer is and, in many cases, know the manufacturer personally. Again, the trademark's long arm function is less reflected in such situations. Thus, when observing trademark data, the researcher should be cautious in not generalizing the findings to the entire economy. External information and knowledge regarding the studied economy's structure and characteristics are crucial in interpreting the data. This is the case of Mandate Palestine. The trademark data reflect, at most, the activity within the modernized segments of the local economy. However, artisan-produced products, typically produced and sold in the same locality, resorts to interpersonal familiarity as an indication of origin and

---

[60] Frank I. Schechter, *The Rational Basis of Trademark Protection,* 40 Harv. L. Rev. 813, 818 (1927).

other product information. The traditional/modernized economy distinction and the transition from the former to the latter may explain the gaps between Jewish-owned and Arab-owned trademarks in Mandate Palestine.

The fourth caveat is a registration bias. Researchers examine available material and may be tempted to assume that trademarks reflect all commercial activity. The previous two caveats pointed to exceptions—open markets and traditional production—where there was no need to register trademarks. But even in other sales modes, such as regular shops and modernized economies, we should remember that registering trademarks was voluntary. Some regulated markets required marking products, such as naming medicines, but registering a trademark with the trademark system was not compulsory.

My study of the emergence of the brand of Jaffa Oranges provides an example.[61] I commenced with studying the trademark data, finding about 100 registered marks relating to citrus. But additional resources—archival material, newspaper reports of the time, memoirs of people in the citrus industry, scholarly literature, and interviews with some family members involved at the time—pointed to *markas*. *Markas* was the name given to citrus marks not registered with the trademark office. Growers and traders had an interest in marking their oranges to differentiate them from local and foreign competitors (mostly Spanish) and to differentiate different qualities of their own products, usually in a triad system: The best oranges were trademarked, the second-class oranges were occasionally trademarked, and the third-class oranges, having the lowest quality, were named, but the traders did not bother to register those names as trademarks. Citrus in the latter category received new names each season, enabling traders to avoid a bad reputation. At some point, there was an official requirement to mark each shipped batch of oranges uniquely. This was a top-down regulation initiated by the British for their own needs: The oranges arrived at their European destinations in a mess. The British insisted on identifying each box and attributing it to its grower for quality assurance, handling, and taxes. This interest commenced as a recommendation in 1927,[62] and, in 1932, became a regulatory duty.[63] However, the citrus regulation said nothing about registering the marks in the trademark registry. The result was new names for the products and a noticeable rise in trademark

---

61 Michael Birnhack, *The Emergence of a Brand: A Case of Jaffa Oranges from Mandate Palestine*, *in* Research Handbook on the History of Trademark Law (Lionel Bently & Robert Bone, eds., forthcoming 2024).

62 *See* Schedule 1 of the Fruit Export Ordinance No. 51 of 1927. The schedule included a form with the intention to export citrus, with an option of naming the brand.

63 *See* Birnhack, *supra* note 61.

applications, but there was also a rise in the number of unregistered *markas*. Thus, the reconstructed registry reveals only part of the story.

Finally, a fifth caveat regarding the use of trademark data concerns its interpretation, especially regarding levels of competition. We saw this in the case of the tobacco industry. The trademark data are a good starting point, as they raise a hypothesis about the level of competition in a specific industry. However, additional anchors are needed to affirm or refute such a hypothesis. Importantly, it matters whether the products were meant for local consumption or for export. If local, the trademark activity may indeed indicate the level of competition. But if the competition was in the foreign markets, there was no reason to register the mark locally. Rather, we would expect the local producer or trader to register their marks in the destination markets, where they would encounter various other barriers. To explore this option, we need access to foreign trademark registrations, but many such historical registries are unavailable. For example, while WIPO has some such registries for former British colonies and mandates, the data do not reveal the applicant's country of origin.[64] Scholars have begun linking trademark registrations using AI tools, but there is still a long road ahead of us.

The five shortcomings and related caveats are not a conclusive list. These are the elements that emerged from the case study of Mandate Palestine, and there may be additional limitations. Researcher, beware!

V. CONCLUDING REMARKS

Historical trademark data can add a new vantage point to explore the practice of law on the ground, thus teaching us about gaps between the law in the books and law in practice. Trademark data may expose trends and patterns, affirming or disproving hypotheses about the intended goals and actual practices. For example, in the context of colonial entities, we may ask for whom the system was deployed. For post-colonial scholars, the immediate answer would be that the system was meant to serve the colonizer, but in the case of Mandate Palestine, we see that it also served the local population, albeit differentially for Jewish traders and Palestinian Arab traders. Trademark data may reflect macroeconomic activity, as well as the dynamics of specific industries, thus raising new questions. The benefits of exploring historical trademark data are substantial, but there are some caveats. One should be cautious in interpreting the data; the data

---

[64] Email correspondence with Ryan Lamb, Statistics and Data Analytics Division, WIPO (Oct. 18, 2022).

reveal much, but not everything. In this regard, supplementing resources is critical before reaching conclusions. Thus, trademark data may triangulate other resources.

The reconstruction of Mandate Palestine's trademark registry offers a few lessons for others who may attempt similar projects: Study the law in advance, plan as much as possible, extract as much as you can, search for abnormalities in the data, code the data, and strive for accuracy and consistency. As for the future, still missing is a global dataset to offer newer insights about the practices of individual business players and the developments of various markets. Digitization and AI tools will likely be useful in reconstructing and analyzing trademark data.

COMMENTARY

INCORPORATING UNCERTAINTY IN TRADEMARK SURVEYS: DO RESPONDENTS REALLY KNOW WHAT THEY ARE TALKING ABOUT?

By Barton Beebe,[*] Roy Germano,[**] Christopher Jon Sprigman,[***] and Joel H. Steckel[****]

I. INTRODUCTION

In this brief commentary, we summarize our recent work related to the current state of trademark surveys.[1] In particular, we identify and present empirical evidence of a fundamental problem with trademark survey evidence: while the leading survey formats in trademark law test for whether consumers hold a particular belief, they do not examine the *strength* or the *varying degrees of certainty* with which consumers hold that belief. Yet, as the social science literature has long recognized, the degree of certainty with which consumers hold particular beliefs shapes their behavior in the marketplace, and thus it should also shape, we believe, how trademark disputes play out in the courtroom.

We demonstrate the relevance of evidence regarding consumer uncertainty in the context of the long-running PRETZEL CRISPS genericness dispute, *Snyder's Lance, Inc. v. Frito-Lay North America, Inc.*[2] Our experiments show that low-cost, easily administered, and relatively simple modifications to common

---

[*] Barton Beebe, John M. Demarais Professor of Intellectual Property Law, New York University School of Law.

[**] Roy Germano, Senior Research Scholar, New York University School of Law.

[***] Christopher Jon Sprigman, Murray and Kathleen Bring Professor of Law, New York University School of Law.

[****] Joel H. Steckel, Professor of Marketing, New York University Stern School of Business.

[1] This commentary is based on an article published by the authors in the *Emory Law Journal*. See Barton Beebe, Roy Germano, Christopher Jon Sprigman, and Joel H. Steckel, Consumer Uncertainty in Trademark Law: An Empirical Investigation, 72 Emory L.J. 489 (2023), *available at* https://scholarlycommons.law.emory.edu/elj/vol72/iss3/1/.

[2] No. 21-1758, 2021 WL 6330712, at *1 (4th Cir. Aug. 31, 2021) (granting a motion to voluntarily dismiss the case). The case originally began under the caption *Frito-Lay N. Am., Inc. v. Princeton Vanguard, LLC*, 109 U.S.P.Q.2d 1949 (T.T.A.B. 2014). Snyder's Lance, Inc. acquired Princeton Vanguard, LLC in 2012.

trademark survey formats can reveal relevant information and provide a richer interpretation of consumer perceptions regarding trademarks.

Our bottom line is clear. Modifying traditional survey formats to elicit evidence regarding consumer uncertainty provides additional information litigants can use to demonstrate, and courts can use to infer, the true state of consumer beliefs about particular trademarks.

II. BACKGROUND

Trademark litigation generally turns on the answer to some version of the following question: What do consumers believe?[3] In trademark infringement cases, the question is whether it is likely that a substantial proportion of consumers mistakenly believe that goods bearing the one party's trademark originate from or have some relationship with another party. Even more fundamentally, for a mark to be protectable in the first place, consumers must believe that the mark refers to a specific producer and not an entire category of goods; i.e., the mark must not be "generic."

Consumer surveys often provide courts with evidence to aid their understanding of what consumers believe. Litigants hire survey experts to survey a sample of a relevant consumer population and then testify about their findings. These surveys can play decisive roles in the outcomes of trademark disputes. Consider the recent closely watched Supreme Court case *United States Patent & Trademark Office v. Booking.com B.V.*[4] The outcome of the case was driven primarily by survey evidence showing that 74.8% of the survey's respondents perceived BOOKING.COM as a brand name.[5] This prompted both Justice Sotomayor in her concurrence and

---

[3] J. Thomas McCarthy, McCarthy on Trademarks and Unfair Competition § 32:158 (5th ed. 2020)).

[4] 140 S. Ct. 2298 (2020).

[5] *Id*. at 2305 ("Consumers do not in fact perceive the term 'Booking.com' [as a generic term], the courts below determined. The PTO no longer disputes that determination. That should resolve this case: Because 'Booking.com' is not a generic name to consumers, it is not generic."). *See also* Booking.com B.V. v. U.S. Pat. & Trademark Off., 915 F.3d 171, 183 (4th Cir. 2019) ("[W]here, as here, the district court found that the survey was methodologically sound, the survey is strong evidence that the public does not understand BOOKING.COM to refer to the proposed mark's generic meaning."). *Cf. Booking.com*, 140 S. Ct. at 2313 (Breyer, J., dissenting) ("What, then, stands in the way of automatic trademark eligibility for every 'generic.com' domain? Much of the time, that determination will turn primarily on survey evidence, just as it did in this case. See 915 F. 3d, at 183–184.").

Justice Breyer in his dissent to warn against placing too much weight on survey evidence in genericism determinations.[6]

In *Snyder's Lance*, a North Carolina federal district court ruled that PRETZEL CRISPS was generic despite survey evidence proffered by the plaintiffs putatively documenting that 55% of respondents stated that PRETZEL CRISPS was a brand name and not a general product category.[7] Here the court's decision was at odds with the survey evidence. Was the survey flawed? Was the court wrong?

In our view, the real lesson from the *Snyder's Lance* litigation is much deeper. That litigation exposes a flaw in the way the legal community and the experts it hires generally go about designing and conducting their trademark surveys. In particular, trademark surveys typically do not include questions surrounding *respondent uncertainty*, i.e., the varying degrees of confidence respondents have in their responses. As currently constituted, the leading survey formats provide no sufficient way for respondents to indicate the strength with which they hold a particular belief.

To illustrate the problem, imagine a pair of household cleansers, AJAX and AJAR. AJAX is a best-selling incumbent brand. In contrast, AJAR is a new market entry; perhaps it has not even appeared in stores yet. Concerned by the similarity of the words AJAX and AJAR, counsel for AJAX files a lawsuit alleging likelihood of confusion between the two brands before the public becomes aware of AJAR. In support of that lawsuit, AJAX proffers a survey in which AJAX and AJAR were shown side by side, likely in the presence of other products, and respondents were asked whether these two products were put out by the same or different companies. Respondents were also given the option of responding "Don't know."

However, since AJAR is unknown to the public, it is hard to imagine that the survey respondent could possibly know with any reasonable degree of certainty whether or not it was put out by the same company as AJAX. The best a respondent could do is provide a subjective belief, given with some degree of uncertainty. As such, the dominant response given to the critical question really should be "Don't know." The other two possibilities (made by the same company or made by different companies) reflect a subjective certainty that is impossible for respondents to have given that they are unfamiliar with AJAR.

While it is true that the common survey formats in theory allow for respondent uncertainty by providing respondents with the option to respond "Don't know," decades of experience show that

---

[6] *Id.* at 2309 (Sotomayor, J., concurring); *id.* at 2313–14 (Breyer, J., dissenting).

[7] Snyder's-Lance, Inc. v. Frito-Lay N. Am., Inc., 542 F. Supp. 3d 371 (W.D.N.C. 2021).

relatively few respondents resort to that response. Social science research has long made clear that survey respondents are typically unwilling to admit, or may be discouraged from admitting, that they don't know or have no opinion.[8] Instead, they engage in what Jon Krosnick has described as "mental coin-flipping" and select answer choices at random.[9] These "nonattitudes" or "pseudo-opinions"[10] look like valid responses, and are treated as such when survey results are aggregated and reported, but they do not measure true underlying attitudes or meaningful beliefs.[11]

Setting the "Don't know" option aside, the remaining response options in trademark surveys present a stark binary choice: the products either do, or do not, originate from the same or different companies (tests of confusion); or the product name designates either a brand or a product category (tests of genericness). In other words, trademark surveys may prompt some respondents to express beliefs they do not actually hold, or which they hold only very weakly. Other respondents may hold multiple conflicting beliefs but are nevertheless forced by the survey format to express just one. Still others, though they may not be guessing or choosing at random, may be uncertain and respond differently if asked the same question at different times—a problem known as "response instability."[12]

This discussion raises a number of straightforward questions relating to trademark surveys. Do trademark survey respondents

---

8 Philip E. Converse, *The Nature of Belief Systems in Mass Publics*, 18 Critical Rev. 1 (1964); *see also* John P. Liefeld, *How Surveys Overestimate the Likelihood of Consumer Confusion*, 93 TMR 939 (2003) (in light of survey respondents' reluctance to state that they have no opinion or do not know, reporting the results of a series of experiments involving trademark surveys using different forms of "filter questions" asking if respondents had a previously formed opinion or attitude available in memory to elicit no opinion or do not know responses).

9 Jon A. Krosnick, *Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys,* 5 Applied Cognitive Psych. 213, 220 (1991).

10 Lee Sigelman & Dan Thomas, *Opinion Leadership & the Crystallization of Nonattitudes: Some Experimental Results*, 16 Polity 484, 484 (1984) ("What is it that prompts as many as one respondent in three to express an opinion, pro or con, on a given issue with absolutely no information or knowledge on the matter to guide his or her response?").

11 English courts are especially sensitive to the problem of nonattitudes in trademark survey evidence and explicitly require surveys to avoid prompting respondents to form beliefs that they would not otherwise have had. According to the "Whitford Guidelines" developed by Mr. Justice Whitford in *Imperial Group plc & Another v. Philip Morris Limited & Another*, [1984] RPC 293, a survey question must not "direct the person answering the question into a field of speculation upon which that person would never have embarked had the question not been put." *Id.* at 303. *See also* Interflora Inc. v. Marks & Spencer Plc, [2013] F.S.R. 21, para. 151 (requiring courts to consider "evidence that any further survey will comply with the Whitford guidelines" when determining whether to grant a party permission to conduct a survey).

12 John Zaller & Stanley Feldman, *A Simple Theory of the Survey Response: Answering Questions Versus Revealing Preferences,* 36 Am. J. Pol. Sci. 579, 580 (1992).

have varying levels of confidence in their survey responses? Do respondents offer responses that take the form of nonattitudes? Are there workable methods by which trademark surveys can test for belief strength and nonattitudes?

The concept of belief strength is fundamental to social science understandings of consumer perception and consumer behavior. Behavioral scientists recommend that "'belief strength,' or more simply 'belief' be measured by a procedure which places the subject along a dimension of subjective probability involving the object (in this case the mark) and some related attribute (in this case its source)."[13] In other words, a respondent's belief, by definition, reflects his or her uncertainty with respect to the object of belief—i.e., the trademark for our purposes. At the same time, the dominant responses to trademark surveys do not reflect that uncertainty. Simply put, unless the respondent says, "Don't know/No opinion," common survey formats imply that the respondent has complete certainty in his or her answer.

That is a real shortcoming with respect to trademark law, because social science evidence strongly suggests that the degree of uncertainty or confidence that a respondent has about the source identification properties of a specific trademark, be they identifying a specific source or the generic nature of the mark, directly impacts how that mark influences real-world consumer purchase decisions.[14] Thus, trademark surveys fail to measure potentially valuable information about respondent uncertainty.

In this commentary, we argue that trademark law, and trademark consumer surveys, in particular, should acknowledge consumer uncertainty—the reality that consumer beliefs are not binary, but held at varying levels of strength. Specifically, we assert that the central inquiry in trademark litigation and trademark consumer surveys should ask: Is it likely that some threshold

---

[13] Martin Fishbein & Icek Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research 12 (1975).

[14] Icek Ajzen, *Consumer Attitudes and Behavior,* in Handbook of Consumer Psychology, 525, 525–48 (Curtis P. Haugtvedt et al. eds., 2008); Jon A. Krosnick & Robert P. Abelson, *The Case for Measuring Attitude Strength in Surveys*, in Questions About Questions: Inquires into the Cognitive Bases of Surveys (Judith M. Tanur ed., 1992); Peter M. Bentler & George Speckart, *Models of Attitude—Behavior Relations*, 86 Psych. Rev. 452, 452–64 (1979); Stephen J. Kraus, *Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature,* 21 Personality and Soc. Psych. Bull. 58 (1995); Denis T. Regan & Russell Fazio, *On the Consistency Between Attitudes and Behavior: Look to the Method of Attitude Formation,* 13 J. Experimental Soc. Psych. 28 (1977); Jaideep Sengupta & Gavan J. Fitzsimons, *The Effect of Analyzing Reasons on the Stability of Brand Attitudes: A Reconciliation of Opposing Predictions*, 31 J. Consumer Rsch. 705, 705–11 (2004); Charles R. Tittle & Richard J. Hill, *Attitude Measurement and Prediction of Behavior: An Evaluation of Conditions and Measurement Techniques*, 30 Sociometry 199 (1967).

proportion of consumers hold a particular belief *at a substantial level of certainty*? To be sure, incorporating consumer belief strength adds an additional degree of complexity to trademark doctrine and fact-finding, but we think that the benefits in information gained far outweigh the costs.

We begin by summarizing a study we conducted based on the *Snyder's Lance* litigation to demonstrate that respondents have greater uncertainty than their survey responses indicate. We also show that the inclusion of a simple uncertainty assessment can change how the results of a trademark survey (in this case a *Teflon* genericness test) are interpreted.[15] We then show that probing for uncertainty in trademark surveys is not in fact entirely new but was a feature of trademark surveys at their origin. Finally, we close with a discussion of the implications of our findings for litigants, survey experts, and courts.

III. INCORPORATING UNCERTAINTY INTO THE *PRETZEL CRISPS* SURVEY

The *Snyder's Lance* dispute arose when snack food behemoth Frito-Lay opposed Snyder's Lance's application to register the term "pretzel crisps" for its pretzel cracker snack. Frito-Lay argued that "pretzel crisps" is a generic term and therefore not registrable. Survey evidence was introduced early in the dispute.[16] Our illustrations build on the *Teflon* survey that Dr. E. Deborah Jay developed as an expert witness for Snyder's Lance.[17]

Teflon surveys are used to determine whether or not an asserted mark is *generic*. A term is generic if most consumers understand it not as indicating the source of any particular product, but rather as denoting a type or category or "genus" of products. So, for example, the term "sugar" is generic for sucrose, whereas the term DOMINO is distinctive for a particular brand of sugar. Terms that function as generic labels, as "sugar" does for sucrose, do not qualify for trademark protection both because consumers do not perceive them

---

15 We present similar empirical work on uncertainty for the *Eveready* and *Squirt* formats for assessing likelihood of confusion in our *Emory Law Journal* article. *See supra* note 1.

16 *See* Princeton Vanguard, LLC v. Frito-Lay N. Am., Inc., 786 F.3d 960 (Fed. Cir. 2015).

17 E. Deborah Jay is principal of Jay Survey Strategics, LLC, where she conducts, evaluates, and testifies about litigation surveys in trademark, deceptive advertising, right of publicity, copyright, patent, wage and hour, and employee discrimination cases, among others. *See* Jay Survey Strategics LLC, http://www.jaysurveystrategics.com (last visited November 20, 2023).

as trademarks and because denying competitors the ability to use such terms would significantly impair competition.[18]

The *Teflon* survey format is generally structured in two parts. The first offers what commentators have described as "essentially a mini-course in the generic versus trademark distinction," followed by a mini-test to confirm that respondents grasp the difference. This mini-test typically runs the respondent through two or three terms (such as "washing machine" and CHEVROLET) to ask whether the terms are common (i.e., generic) names or brand names. After respondents have proven that they understand the difference, the second part of the *Teflon* survey then presents respondents with six or seven terms, including the mark at issue, in this case PRETZEL CRISPS, and asks respondents to classify each as either a brand name or a generic term. Importantly, respondents are forced to choose among only three possible answers: "Brand name," "Generic name," or "Don't know."[19] Dr. Jay reported that 55% of her respondents classified PRETZEL CRISPS as a brand name, a statistic that supported, though not strongly, her client's case. Though the second part of the *Teflon* survey format allows for a "don't know" option, it fails to capture important information about a respondent's degree of uncertainty. In Dr. Jay's study, only 9% of respondents chose the "Don't know" option—a figure that we believe to be unreasonably low.[20]

To examine this issue, we exposed 242 respondents to a *Teflon* survey modeled after the survey Dr. Jay administered for Snyder's Lance. In a departure from Dr. Jay's methodology, we randomly assigned respondents to one of three groups.[21] Each group saw a

---

[18] U.S. Pat. & Trademark Off. v. Booking.com B.V, 140 S. Ct. 2298, 2301 (2020) ("A generic name—the name of a class of products or services—is ineligible for federal trademark registration.").

[19] For example, the main survey in the *Teflon* case read:

> I'd like to read 8 names to you and get you to tell me whether you think it is a brand name or a common name; by *brand* name, I mean a word like *Chevrolet* which is made by one company; by *common* name, I mean *a word like automobile* which is made by a number of different companies. So if I were to ask you, "Is Chevrolet a brand name or a common name?," what would you say? Now, if I were to ask you, "Is washing machine a brand name or a common name?," what would you say?

McCarthy, *supra* note 3, § 12:16.

[20] Snyder's Lance, Inc. v. Frito-Lay N. Am., Inc., 542 F. Supp. 3d 371, 399 (W.D.N.C. 2021).

[21] We recruited subjects for these studies through Amazon Mechanical Turk (MTurk). MTurk is a crowdsourcing service that allows researchers to recruit large numbers of participants for online studies. People who responded to our call for subjects on MTurk were directed to an online survey we created in Qualtrics. For our *Teflon* study, 392 people responded to our call for subjects on MTurk. Following Dr. Jay's procedures, we determined whether people were eligible to participate in the study by asking the following questions: "In the past 3 months, did you, personally, purchase salty snacks for you or someone else?"; and (2) "In the next 3 months, do you think you, personally, will

different variation on the standard *Teflon* format. The purpose of these variations was to measure how certain the survey respondents were about whether a term is generic or represents a brand name. We explain one variation here, a forced-choice question with follow-up, and we refer the reader to our *Emory Law Journal* article for more detailed descriptions of the implementation and results of investigating the other two. We do note, however, that the results of the three approaches examined there are consistent and all lead to the same qualitative conclusions.

IV. USING A FORCED-CHOICE QUESTION WITH FOLLOW-UP

Of our 242 respondents, 81 were assigned to a forced-choice question with a follow-up question. These 81 respondents proceeded through a survey nearly identical in format to Dr. Jay's, except that we added an additional follow-up question intended to probe respondents' degree of confidence in their classification of the term PRETZEL CRISPS.

Specifically, after passing the same mini-test that Dr. Jay used, these respondents were exposed in random order to the same six control terms that Dr. Jay used: CHEESE NIPS, MACADAMIA NUTS, ONION RINGS, GOURMET POPCORN, FLAVOR TWISTS, and SUN CHIPS. Also, as Dr. Jay did, we then exposed respondents to the term PRETZEL CRISPS. Upon exposing respondents to each term, we asked: "Do you think [the term] is a brand name or a generic name?" Below the question, respondents saw three answer choices: "Generic name," "Brand name," and "Don't know/Not sure."[22] After this question, we added a follow-up question that Dr.

---

purchase salty snacks for you or someone else?" Anyone who did not answer in the affirmative to one of these questions was considered ineligible. Next, potential subjects read a set of instructions that explained the difference between a brand name and a generic name, followed by two practice questions to make sure they understood this difference. The practice questions asked subjects whether the terms BAKED TOSTITOS and TORTILLA CHIPS are brand names or generic names. Only people who answered both practice questions correctly (the first is a brand, the second generic) were permitted to participate in the study. Of the 392 people who responded to our call for subjects on MTurk, 242 were deemed eligible to participate in the full survey. These 242 subjects were then randomly assigned to one of three groups. On the reliability of MTurk, see Barton Beebe, Roy Germano, Christopher Jon Sprigman, and Joel H. Steckel, *Testing for Trademark Dilution in the Court and the Lab*, 86 U. Chi. L. Rev. 611, 662–65 (2019). *See also* Matthew J.C. Crump, John V. McDonnell, and Todd M. Gureckis, *Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research*, 8 PLOS ONE 1, 3–11 (2013), *available at* https://doi.org/10.1371/journal.pone.0057410 (last visited Dec. 6, 2023).

[22] We randomized the order in which the words "brand name" and "generic name" appeared in the questions and answer choices.

Jay did not ask. As Figure 1 below shows, respondents who answered "Generic name" or "Brand name" were asked: "How likely do you think it is that your answer is correct?" The answer choices were arrayed horizontally and included, "Just guessing," "Somewhat likely correct," "Very likely correct," and "Definitely correct." Respondents who answered "Don't know" to the first question were not asked the follow-up question. This follow-up question served to assess respondent uncertainty.

Figure 1
Follow-up Question Format

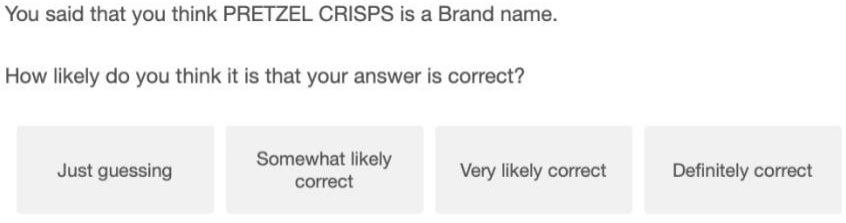
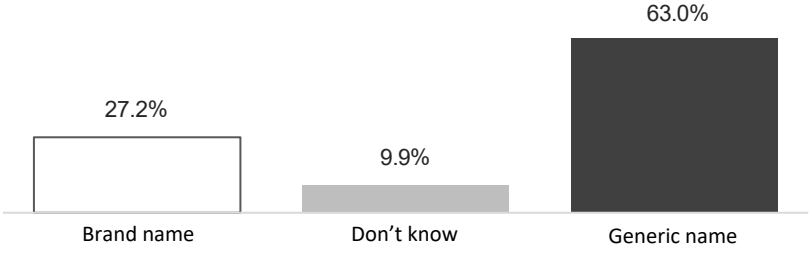


Figure 2 reports the distribution of responses to the first PRETZEL CRISPS question. A clear majority of respondents, 63%, indicated that they believed PRETZEL CRISPS to be a generic term, while just 27% perceived PRETZEL CRISPS as a brand name. Consistent with our experience in other *Teflon* studies, a relatively small percentage, 9.9%, responded "Don't know/Not sure." Taken at face value, these results support a finding that the term is generic, a result at variance with Dr. Jay's conclusion, even though the percentage of "Don't know/Not sure" responses was approximately the same as hers.

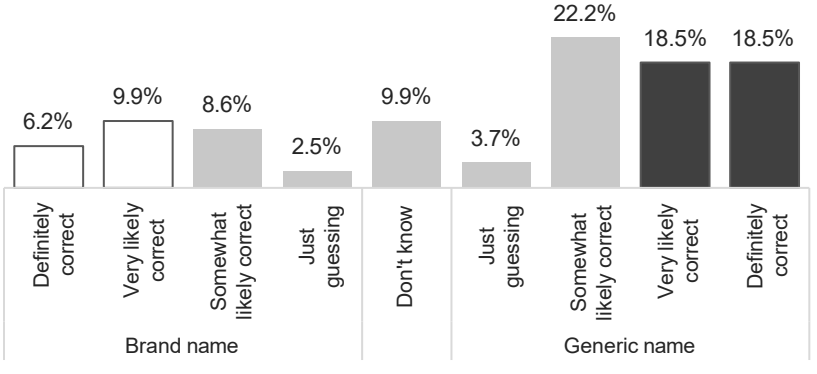
Figure 2
Distribution of Group A Responses to the Standard *Teflon* Question



The distribution of responses to the follow-up question, however, reveals substantial uncertainty hidden behind respondents'

answers to the first question. Figure 3 reports the percentage of respondents who gave each possible combination of answers to the first PRETZEL CRISPS question and the follow-up question. For example, the left-most bar labeled "Brand name: Definitely correct" refers to the percentage of respondents, 6.2%, who answered "Brand name" to the first question and "Definitely correct" to the follow-up question. The center bar represents the 9.9% of respondents who answered "Don't know" to the first question and who were thus not asked the follow-up question.

Figure 3
Distribution of Responses to Both Questions



We observe in Figure 3 that nearly half of respondents, 46.9%, expressed high levels of uncertainty about their answer to the first PRETZEL CRISPS question.[23] The five middle bars in Figure 3, shaded in light gray, represent these uncertain respondents. Remarkably, 6.2% of respondents admitted in their answer to the follow-up question that they were just guessing in their answer to the first question. An additional 30.8% of respondents indicated that they believed their answer to the first question to be only "somewhat likely correct." Put differently, of those respondents who answered "brand name" to the first question, 59% responded "definitely correct" or "very likely correct" to the follow-up question, and of

---

[23] From the perspective of the mechanics of trademark litigation, the legal standard in trademark cases (as in virtually all civil litigation) is preponderance of the evidence—i.e., that consumer confusion is more likely than not. And once we redesign trademark surveys to take belief strength into account, it is only the top two points on the Likert scale that represent a belief arguably strong enough to indicate that the proposition is more likely true than not. Beliefs of this strength, moreover, are more likely to impact consumer behavior. For these reasons, the party bearing the burden of persuasion on the question addressed by a survey should not be able to rely on consumers who admit to guessing or who report that they are merely somewhat likely to perceive the mark or marks at issue in the manner alleged.

those who answered "generic name" to the first question, 58% answered "definitely correct" or "very likely correct." The remaining respondents who answered "brand name" or "generic name" were guessing or less convinced of the belief they expressed. Whichever way the data are parsed, it is clear that the first PRETZEL CRISPS question on its own provides limited information about the reality of consumer beliefs with respect to the term, and at worst, possibly misleading information about those beliefs.

The widespread uncertainty that respondents reported may help to explain why our results on the first PRETZEL CRISPS question differed from Dr. Jay's. We used the same methodology and a sample qualified using the same mini-test questions.[24] Reassuringly, our results and hers were virtually identical on the six control terms, five of which presented a relatively easy case about which respondents were more likely to have strongly held beliefs.[25] Yet with respect to PRETZEL CRISPS, while we found that only 27% of our respondents indicated that they perceived PRETZEL CRISPS as a brand name, Dr. Jay reported that 55% of her respondents did so, a statistic that supported, though not strongly, her client's case.

In essence, our study and Dr. Jay's came to opposite conclusions about the main fact question the survey was intended to resolve. But as we noted, when we look behind our results on the first PRETZEL CRISPS question, we find that many of the respondents held weak beliefs or nonattitudes. We suspect that had Dr. Jay asked our follow-up question, she would have discovered that the same was true of a high proportion of individuals in her sample. We cannot prove that weakly held beliefs and nonattitudes are responsible for the differences between our results and Dr. Jay's on the first PRETZEL CRISPS question. It is revealing, however, that our results were so similar to Dr. Jay's on the control terms, which

---

[24] *See supra* Part I.B. In *Frito-Lay N. Am., Inc. v. Princeton Vanguard, LLC*, 124 U.S.P.Q.2d 1184 (T.T.A.B. 2017), the Trademark Trial and Appeal Board criticized Dr. Jay's use in her survey's initial mini-course of the term WHEAT THINS as an example of a brand, since it is a highly descriptive mark "and thus not a good example to participants of how to distinguish between a distinctive term and a merely well-advertised highly descriptive or even generic term." *Id.* at 1197. To replicate Dr. Jay's protocol, we used the same example. Thus, we cannot point to the circumstances of the mini-course to explain the difference between Dr. Jay's and our results.

[25] With regard to the three brand names, 96% of Jay's sample and 98% of our sample said that SUN CHIPS is a brand name; 85% of Jay's sample and 89% of our sample said that CHEESE NIPS is a brand name; and 48% of Jay's sample and 52% of our sample said that FLAVOR TWISTS is a brand name. With regard to the three generic names, 92% of Jay's sample and 96% of our sample said that MACADAMIA NUTS is a generic name; 91% of Jay's sample and 98% of our sample said that ONION RINGS is a generic name; and 72% of Jay's sample and 86% of our sample said that GOURMET POPCORN is a generic name.

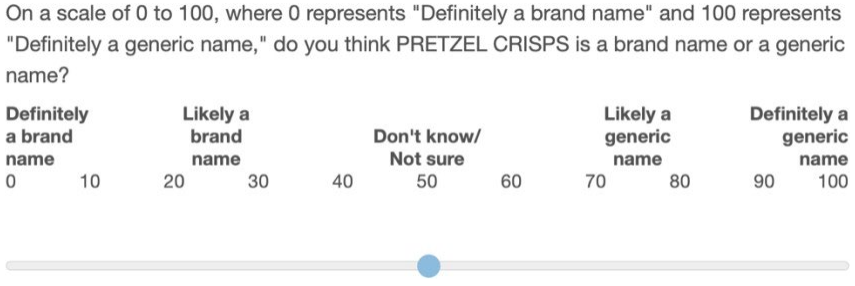
were designed, on the whole, to fall more clearly in the category of either brand name (e.g., CHEESE NIPS) or generic term (e.g., MACADAMIA NUTS). Respondents in both studies largely agreed on those terms. They only diverged on the more difficult question of whether PRETZEL CRISPS is generic or a brand name.

In our *Emory Law Journal* article, you will find discussion of two other ways of assessing respondent uncertainty. In the first, eighty-one respondents were not provided with the "Generic name"/"Brand name"/"Don't know" answer choices, but rather with a seven-point Likert scale as shown in Figure 4. This answer format allowed respondents to state whether they believe the term is a brand name or a generic term and simultaneously signal their level of confidence in that belief.

Figure 4
Question Structure: Likert Scale

Finally, a separate group of eighty respondents was provided with a slider that ranged from 0 to 100, where 0 represented "Definitely a brand name" and 100 represented "Definitely a generic name." The slider presented respondents with a continuous scale, but we placed intermediate labels on the slider between the 20 and 30 marks ("Likely a brand name"), at the 50 mark ("Don't know/Not sure"), and between the 70 and 80 marks ("Likely a generic name") to help guide the respondent. Figure 5 shows the format of the slider presented to these respondents.

Figure 5
Question Structure: Slider



These formats differ, and the results of the tests using them differ around the edges. For reasons we explain in our *Emory Law Journal* article, we advise use of the Likert scale rather than the sliding scale.[26] Regardless, the main conclusions remain the same. We believe that responses of "somewhat likely correct," "just guessing," or "don't know" reflect uncertainty on the Likert scale, as do responses of 15 to 85 on the slider. Approximately half the respondents exhibited substantial uncertainty about whether PRETZEL CRISPS was a brand name or a generic term (46.9% for the two-stage forced-choice questioning, 44.5% for the Likert scale, and 47.5% for the slider), and approximately a quarter or less believed it was a brand name with any reasonable degree of certainty (16.1% for the two-stage forced-choice questioning, 21.0% for the Likert scale, and 12.5% for the slider). Indeed, these results convey a very different meaning than Dr. Jay's on the main question of whether consumers perceive PRETZEL CRISPS as being from a specific producer or representing a product category. Accounting for uncertainty refines the conclusions of a study and has the potential to change them entirely. Our results support the court's decision that the PRETZEL CRISPS mark was generic, despite what we believe is Dr. Jay's largely proper application of the standard *Teflon* methodology as it currently stands.

V. LEARNING FROM HISTORY

The procedures discussed above, along with those in our *Emory Law Journal* article for the *Eveready* and *Squirt* survey approaches, detail low-cost, easily administered, and relatively simple modifications to the standard formats of trademark surveys that, by registering consumer uncertainty, will provide courts with what we believe is substantially better information about consumer beliefs. However, as much as we would like to take credit for the simple, compelling idea of incorporating uncertainty into trademark surveys, we cannot.

It is unfortunate that the early history of trademark survey evidence has been largely forgotten because there is much we can learn—or re-learn—from it. The trademark survey formats first proposed by social scientists a century ago actually examined consumer belief strength. The story of trademark survey evidence over the past century is a story of regression to the blunt instruments used today.

The first survey evidence ever submitted to an American court in a trademark dispute was introduced by the Coca-Cola Company

---

[26] *See supra*, n.1, at 522–23.

in 1921.[27] In 1915, Coca-Cola opposed the registration of the mark CHERO-COLA for cola-flavored soft drinks, asserting that it was confusingly similar to the mark COCA-COLA.[28] Coca-Cola submitted a survey report as evidence that consumers would confuse the marks. In the report, the Columbia University–affiliated psychologist Richard Paynter described four experiments he conducted under laboratory conditions.[29] In the first of these experiments, the respondent was shown in random order twenty slips of paper on each of which was typed a word mark.[30] One of the slips of paper bore the mark "Coca-Cola."[31] After a brief pause, the respondent was then shown in random order forty slips of paper, twenty of which bore marks not previously presented to the respondent, nineteen of which bore marks previously shown to the respondent, and one of which bore the mark "Chero-Cola" instead of "Coca-Cola."[32] For these forty slips of paper, the written instructions provided to each respondent explained:

> [Y]ou will be . . . asked to pick out those marks you have just seen in the presentation and those which you have not seen. You will be further asked to sort the marks into six piles, according to the degree of your confidence or certainty of your recognition of your marks. There are three degrees of certainty for the marks that are recognized as seen, and three similar degrees for those that are recognized as not seen. The three degrees are "absolutely certain," "reasonably certain," and a "faint idea."[33]

---

27 *See* Coca-Cola Co. v. Chero-Cola Co., 273 Fed. 755 (D.C. Cir. 1921).

28 *See* Edward S. Rogers, *An Account of Some Psychological Experiments on the Subject of Trade-Mark Infringement*, 18 Mich. L. Rev. 75, 77 (1919).

29 *See id.* 77–99 (reproducing the report). *See also* Richard H. Paynter, *A Psychological Study of Trade-Mark Infringement* 42 Archives Psych. 1 (1920) (discussing experiments related to trademark infringement); Richard H. Paynter, *A Psychological Study of Confusion Between Word Trade-Marks*, 11 Bull. U.S. Trade-Mark Assoc. 101 (1915) (reporting the results of experiments similar to those Paynter used in *Coca-Cola Co. v. Chero-Cola Co.* to determine the likelihood of confusion between the words KREMENTZ and KREMO as applied to collar buttons). Paynter's experiments were similar to those envisioned, but not carried out, by Hugo Münsterberg. *See* Hugo Münsterberg, Psychology and Industrial Efficiency 282–293 (1913).

30 *See* Rogers, *supra* note 25 at 78–79. The slips were presented in random order except that "Coca-Cola" and "Chero-Cola" appeared neither first nor last.

31 *Id.* at 79.

32 *Id.* The marks were capitalized but not typed in all uppercase characters.

33 *Id.* at 80. Paynter's second and third experiments followed a similar protocol, except that the second experiment included on each slip of paper below the mark the product category for which the mark was used (e.g., "Soft Drink") and the third experiment included, for purposes of comparison, marks and their products from various recently litigated trademark cases. *Id.* at 79. The fourth experiment exposed the respondents to ten pairs of marks that were the subject of recent trademark infringement cases and

Soon after the *Chero-Cola* case, a trademark litigant once again submitted survey evidence in which respondents were asked to specify their degree of certainty. This time the survey expert was Harold Burtt, an Ohio State University–affiliated psychologist.[34] Burtt roughly followed Paynter's protocols but used a seven-point scale of certainty.[35]

In subsequent decades, trademark survey methods shifted primarily to face-to-face interviews with consumers, conducted either door-to-door or by intercepting consumers in or outside stores. None of these interview-based surveys appears to have probed respondents for their degree of certainty in their response, perhaps because in most cases the trademark owner was the party who submitted the survey and would not likely have benefitted from data showing respondent uncertainty.[36]

That said, from time to time, courts picked up on the problem. Indeed, in one of the most influential judicial analyses of trademark survey evidence in the midcentury, the court in *General Motors Corp. v. Cadillac Marine & Boat Co.*[37] criticized the plaintiff's survey for failing to "take into consideration the hazy and qualified answers" of the survey's respondents.[38] General Motors produced automobiles under the mark CADILLAC; Cadillac Marine & Boat produced boats under the same mark.[39] In an early version of what became the *Eveready* survey format, General Motors asked "Who do you think puts out the boat shown on the opposite pages?"[40] and

---

asked the respondents to order the pairs according to the degree of confusion that the respondent believed each pair would create in consumers. *Id.* at 91–98.

34 Harold E. Burtt, *Measurement of Confusion Between Similar Trade Names*, 19 Ill. L. Rev. 320 (1924).

35 *Id.* at 325–26.

36 *See, e.g.*, Lerner Stores Corp. v. Lerner, 162 F.2d 160, 162 (9th Cir. 1947) (discussing an intercept survey conducted in front of the plaintiff's store); du Pont Cellophane Co. v. Waxed Prods. Co., 6 F. Supp. 859, 878 (E.D.N.Y. 1934) (discussing a house-to-house survey testing whether respondents perceived cellophane as a generic term); Oneida, Ltd. v. Nat'l Silver Co., 25 N.Y.S. 2d 271, 287–88 (N.Y. Sup. Ct. 1940) (discussing two house-to-house surveys). *See also* Beverly W. Pattishall, *Reaction Test Evidence in Trade Identity Cases*, 49 TMR 145, 156 (1959) (arguing that fixed form interviews are the best means of testing for likelihood of consumer confusion); Robert C. Sorensen & Theodore C. Sorensen, *The Admissibility and Use of Opinion Research Evidence*, 28 N.Y.U. L. Rev. 1213, 1215–16 (1953) (arguing that personal interviews are most useful technique of determining public opinion); Robert Bonynge, *Trademark Surveys and Techniques and Their Use in Litigation*, 48 ABA J. 329 (1962) (reviewing mid-twentieth century trademark survey methods).

37 226 F. Supp. 716 (W.D. Mich. 1964).

38 *Id.* at 736.

39 *Id.* at 719–20.

40 *Id.* at 734 n.16.

"Will you please name anything else put out by the same concern?"[41] The court closely scrutinized the survey respondents' answers and repeatedly noted respondents' "unclear or ambiguous" responses.[42] One of the examples the court gave was the response "'Well, since it says "Cadillac," I guess it's Cadillac.'"[43] The survey's questions had not probed for uncertainty, but the respondents' verbatim responses revealed it anyway.

Ultimately, the court in *Cadillac Marine* rejected the plaintiff's efforts to do what we believe so many current surveys seek to do, which is hide respondent uncertainty behind bottom-line, summary percentages of those confused and not confused. "Such qualified answers," the court explained, "are not susceptible to a categorization such as plaintiff attempted in summarizing the poll."[44] Other courts of the time were similarly critical of trademark survey evidence,[45] and even as late as the early 1970s, courts remained generally hostile to it.[46] Things changed with the Seventh Circuit's 1976 opinion in *Union Carbide Corp. v. Ever-Ready Inc.*,[47] which largely inaugurated the current era in which survey evidence plays a substantial role in trademark litigation.

---

41 *Id.*

42 *Id.* at 735.

43 *Id.*

44 *Id.* The *Cadillac Marine* court further criticized the second main question as leading. In the court's view, it prompted respondents who "drew a complete blank," *id.* at 736, on the first question eventually to think of General Motors: "One individual said, 'I have no idea,' in answer to the first question yet the second question brought the answer, 'car.'" *Id.*

45 *See, e.g.*, Nat'l Biscuit Co. v. Princeton Mining Co., Inc., 137 U.S.P.Q. 250 (T.T.A.B. Feb. 12, 1963). In rejecting the plaintiff's survey, the court noted that a review of the survey sheets from which the summary was prepared discloses that the figure in question includes many persons who named opposer or its products only after prefacing their answers with such statements as "I have no idea," "I haven't the slightest idea," "You've got me," "Well golly, I don't know," and the like. *Id.* at 252.

46 *See, e.g.*, Am. Basketball Ass'n v. AMF Voit, Inc., 358 F. Supp. 981, 986 (S.D.N.Y. 1973) (assessing the plaintiff's secondary meaning survey as unworthy of "any substantial weight"); Sears, Roebuck & Co. v. Allstate Driving Sch., Inc., 301 F. Supp. 4, 18 (E.D.N.Y. 1969) (noting that "[o]ne of the dangers inherent in a consumer reaction test is that it is not administered in the context of the market place. Respondents to such a test do not consider those factors which are relevant to the particular purchasing decision at hand."); Aerojet-Gen. Corp. v. Cincinnati Screen Process Supplies, Inc., 172 U.S.P.Q. 114, 118 (S.D. Ohio 1971) (assessing the defendant's likelihood of confusion survey as entitled to "very little weight").

47 531 F.2d 366 (7th Cir. 1976).

VI. CONCLUDING REMARKS

Given the decisive role that consumer perceptions play in the outcome of trademark disputes, it is of the utmost importance that courts understand what consumers actually believe. To do so, courts typically address their analysis to *populations* of relevant consumers and assess those populations probabilistically. The "likelihood of confusion" cause of action prompts courts to ask whether the defendant's trademark is likely to cause the relevant population of consumers to mistakenly believe that the plaintiff's and defendant's goods originate from the same source.[48] The antecedent question of distinctiveness—i.e., whether an asserted mark is protectable at all—is implicitly framed in the same way: to establish the distinctiveness of a descriptive term[49] or an element of product design trade dress,[50] courts consider how likely it is that a substantial proportion of the relevant consumer population perceives the term or element as distinctive of source.

Importantly, in assessing consumer beliefs, trademark law recognizes that most populations of relevant consumers are not homogenous. That is why the likelihood of confusion cause of action does not require courts to find that it is likely that the *entire* population of relevant consumers is confused. Instead, trademark law asks courts to look inside the population of relevant consumers and determine whether an appreciable proportion of that population (typically, 20% to 25%,[51] but sometimes as low as 15% or even

---

48 For federally registered marks, Section 32 of the Lanham Act brands a defendant's use as actionable trademark confusion if it is "likely to cause confusion, or to cause mistake, or to deceive." Lanham Act § 32, 15 U.S.C. § 1114(1). For unregistered marks, Section 43(a) of the Lanham Act defines an infringing use as one "likely to cause confusion, or to cause mistake, or to deceive as to the affiliation, connection, or association" of the junior user with the senior user. Lanham Act § 43(a), 15 U.S.C. § 1125(a).

49 Descriptive terms are protectable as marks if the plaintiff establishes that they have acquired distinctiveness (sometimes referred to as "secondary meaning"); i.e., that an appreciable number of consumers perceive them as indicating the source of particular products or services. Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 163 (1995). A class of "inherently distinctive" marks—i.e., fanciful, arbitrary, and suggestive marks—are protected without the need for plaintiff to establish distinctiveness. Abercrombie & Fitch Co. v. Hunting World, Inc., 537 F.2d 4, 9–10 (2d Cir. 1976).

50 *See* Wal-Mart Stores, Inc. v. Samara Bros., 529 U.S. 205, 216 (2000) (holding that "in an action for infringement of unregistered trade dress under § 43(a) of the Lanham Act, a product's design is distinctive, and therefore protectible, only upon a showing of secondary meaning").

51 *See, e.g.*, McDonald's Corp. v. McBagel's, Inc., 649 F. Supp. 1268, 1277–78 (S.D.N.Y. 1986) (25% supports finding of likely confusion); Bell v. Starbucks U.S. Brands Corp., 389 F. Supp. 2d 766, 776 (S.D. Tex. 2005), *judgment aff'd*, 205 Fed. Appx. 289 (5th Cir. 2006) (25% is sufficient to show a "significant" level of actual confusion and to support a finding of infringement); *see also* McCarthy, *supra* note 3, § 32:188 ("Generally, figures in the range of 25% to 50% have been viewed as solid support for a finding of a likelihood of confusion. In the author's view, survey confusion numbers that go below 20% need to

lower[52]) is confused. If a methodologically sound survey shows that the defendant's conduct will confuse more than that threshold proportion, then a court should find infringement. Trademark law takes the same approach when inquiring whether a particular designation functions as a mark in the first place—although the threshold is typically set higher.[53] As with likelihood of confusion, tests for trademark genericism and distinctiveness base their findings on the percentage of the relevant consumer population that perceives the indicium at issue as indicating source.[54]

That said, although trademark law recognizes the heterogeneity of beliefs within a given population of consumers, the empirical sophistication of trademark law stops there. It does not go deeper to consider the strength and meaningfulness of the beliefs held by each individual within that population. For example, while trademark law's likelihood of confusion analysis assesses consumer populations in continuous terms as proportionally more or less confused, it typically assesses individuals within those populations as binaries; each is either absolutely confused or absolutely not confused. The same is true for trademark law's distinctiveness analysis. It assesses consumer populations in continuous terms as manifesting a proportionally higher or lower incidence of belief that a particular asserted mark indicates the source of a product. But it treats individuals within those populations as binaries: each individual either totally supports or totally rejects the proposition that the asserted mark indicates that source.[55]

At the foundation of trademark surveys is thus an unrealistic simplification of the individual beliefs that, in the aggregate, determine protectability and liability in trademark cases.

---

be carefully viewed against the background of other evidence weighing for and against a conclusion of likely confusion.").

52 Exxon Corp. v. Tex. Motor Exch., Inc., 628 F.2d 500 (5th Cir. 1980) (survey showing 15% confusion was "strong evidence" of a likelihood of confusion where other evidence was also strongly supportive). *See also* McCarthy, *supra* note 3, § 32:188 (reviewing case law relying on a 15% rate of confusion in survey evidence as probative of likely confusion).

53 *See, e.g.*, Spraying Sys. Co. v. Delavan Inc., 975 F.2d 387, 394 (7th Cir. 1992) ("While a 50-percent figure is regarded as clearly sufficient to establish secondary meaning, a figure in the thirties can only be considered marginal.").

54 *See* Wal-Mart Stores, Inc. v. Samara Bros., Inc., 529 U.S. 205, 211 (2000) ("[A] mark has acquired distinctiveness, even if it is not inherently distinctive, if it has developed secondary meaning, which occurs when, 'in the minds of the public, the primary significance of a [mark] is to identify the source of the product rather than the product itself.' Inwood Laboratories, Inc. v. Ives Laboratories, Inc., 456 U.S. 844, 851, n. 11 (1982).").

55 *See* Itamar Simonson, *Trademark Infringement from the Buyer Perspective: Conceptual Analysis and Measurement Implications*, 13 J. Pub. Pol'y & Mktg. 181, 195 (1994) (noting that trademark surveys typically fail to account for respondents' degree of confidence in their responses).

Trademark surveys typically treat an individual consumer's subjective probability as either 1 or 0, then add up the 1s in a relevant population, and from that derive a proportion of consumers who either do or do not hold a particular belief. In contrast, the social science literature has long recognized the obvious: individual beliefs are not binaries.

Our experiments reveal that consumers experience varying degrees of uncertainty in assessing whether a mark is generic or distinctive, or whether two similarly branded products originate from the same source. Current standard trademark survey formats fail to register these degrees of uncertainty. There is a substantial danger, in other words, that trademark surveys may prompt some respondents to provide responses unrelated to their actual marketplace beliefs, if they have any, on a particular question. Other respondents may hold multiple conflicting beliefs but are nevertheless forced to express just one. Still others, though they may not be guessing or choosing at random, may be uncertain and produce significant "response instability"[56]—if asked the same question at a later time, they may respond differently.

Response uncertainty and instability present several implications for litigants, survey experts, and courts. First, survey experts should no longer be given license to hide the reality of respondent uncertainty from the finder of fact. Second, courts should take into account the strength of consumer beliefs when determining whether the plaintiff has satisfied its burden of persuasion. Evidence showing, for example, that some proportion of consumers believe it to be only *somewhat likely* that the defendant's mark originates from the plaintiff should not be the basis for trademark liability, not least because such a weakly held belief may be dispelled when consumers are making decisions in an actual market setting, which almost always provides context that the survey environment lacks. Allowing a plaintiff to use weakly held beliefs to satisfy its burden may be especially inappropriate if a larger share of the population of consumers appears to hold the opposite belief with greater certainty. Third, when a court does find liability, it should consider the strength of consumers' mistaken confusion as to source in tailoring an appropriate remedy. Remedies short of an outright injunction, such as modifications to the defendant's mark or requiring a disclaimer, may be just as effective in disabusing consumers of their weakly held, mistaken beliefs while at the same time limiting the costs imposed on plaintiffs' competitors. Evidence of weakness of survey respondents' beliefs might provide grist for a defendant to overcome the recently codified

---

[56] John Zaller & Stanley Feldman, *A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences,* 36 Am. J. Pol. Sci. 579, 580 (1992).

rebuttable presumption of irreparable harm, for purposes of arguing that no injunction should issue.[57]

Practical questions remain to be resolved. If finders of fact in trademark litigation should take into account consumer uncertainty, then how exactly should they do so? We have made a first step in that direction. Our view is that small degrees of belief strength do not satisfy a plaintiff's prima facie case, meet its burden of persuasion, or justify a blanket injunction of the defendant's conduct. Among the excellent comments we received from anonymous reviewers of this commentary, one asked why litigants would voluntarily design surveys to take uncertainty into account when doing so risks significantly weakening the strength of that evidence. Why indeed? At the least, our hope is that rebuttal reports that challenge a survey expert's findings for failure to account for strength of respondents' beliefs or degrees of uncertainty will provide an opportunity to educate the finder of fact and allow for these principles to be adopted by the courts. At best, the opposing party may have the resources to run its own survey that tests for consumer belief strength and uncertainty.

Finally, a more nuanced understanding of consumer uncertainty in the marketplace may allow for more nuanced forms of relief. With respect to injunctive relief, all the major fields of intellectual property law have grown increasingly sensitive in recent decades to the need for courts to fashion more tailored injunctions.[58] Though trademark law has shared in this trend, it significantly lags behind patent and copyright law. One reason for this may be that trademark courts are not provided with sufficient information about marketplace realities. Current trademark surveys contribute to this problem by giving the impression that there either is or is not confusion and that a term is either a brand name or a generic term. We assert that this oversimplifies what consumers believe. As courts become aware of the wide diversity of consumer beliefs, they may become more comfortable with forms of injunctive relief that fall short of outright prohibitions. Survey evidence that indicates primarily that respondents were only somewhat likely to be confused should not support a finding of actual confusion.

In sum, our experimental findings show that current survey methods deprive courts of information useful in designing effective and appropriately tailored remedies in cases where plaintiffs do prevail. Properly designed trademark surveys would provide courts with the information they need to take account of consumer belief

---

[57] Trademark Modernization Act of 2020, Pub. L. No. 116-260, 134 Stat. 2200 (2020), *codified at* 15 U.S.C. § 1116(a).

[58] *See* Aurelia Hepburn-Briscoe, *Irreparable Harm in Patent, Copyright, and Trademark Cases After* eBay v. Mercexchange, 55 Howard L.J. 643 (2012).

strength when designing remedies. For example, when a survey reveals that an appreciable number of consumers believe mistakenly that the parties' products come from the same source, but for many that belief is only weakly held, then it may be unnecessary for a court to issue an absolute prohibition on the defendant's accused mark. Disclaimers or modest changes in the defendant's mark may be sufficient to disabuse consumers of weakly held mistaken beliefs. And such tailored relief may avoid imposing unneeded costs both on the plaintiff mark-owner's good faith competitors and—perhaps most importantly—on consumers who are *not* confused.

GUIDELINES FOR SUBMISSIONS TO *THE TRADEMARK REPORTER*

Founded in 1911, *The Trademark Reporter* (TMR) delivers engaging and comprehensive peer-reviewed scholarship on trademarks and complementary intellectual property from practitioners, academics, and judges worldwide.

The TMR welcomes submissions of articles, commentaries, and book reviews at wknox@inta.org.

The complete TMR Submission Guidelines are available at https://www.inta.org/resources/the-trademark-reporter/tmr-submission-guidelines/.

THE TMR ONLINE

International Trademark Association: Issues of the TMR beginning with Volume 105, Number 1, 2015, are available to the public at https://www.inta.org/TMR; INTA members may log in for complimentary access to the complete archive.

HeinOnline: Issues of the TMR beginning with Volume 1, Number 1, 1911, are available through HeinOnline.

LexisNexis: Issues of the TMR beginning with Volume 31, Number 1, 1941, are available through LexisNexis.

Westlaw: Issues of the TMR beginning with Volume 80, Number 1, 1990, are available through Westlaw.

PRINTED COPIES

Printed copies of the TMR can be ordered from William S. Hein & Co., Inc. (order@wshein.com).