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 TMR (ISSN 0041-056X) is published electronically six times a year by INTA, 675 Third Avenue, New York, NY 10017-5704 USA. INTA, the INTA logo, INTERNATIONAL TRADEMARK ASSOCIATION, THE TRADEMARK REPORTER, and inta.org are trademarks, service marks, and/or registered trademarks of the International Trademark Association in the United States and certain other jurisdictions.

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.

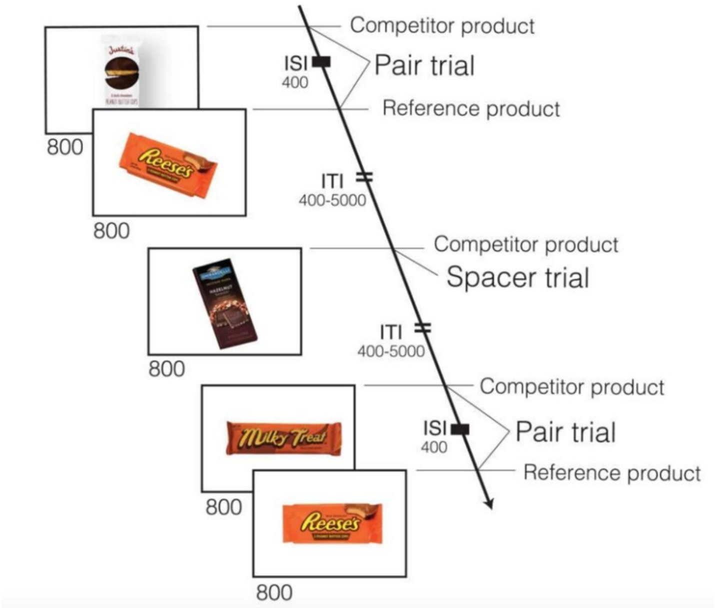

Figure II. Example trial timeline of the fMRI experiment.[49]

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).

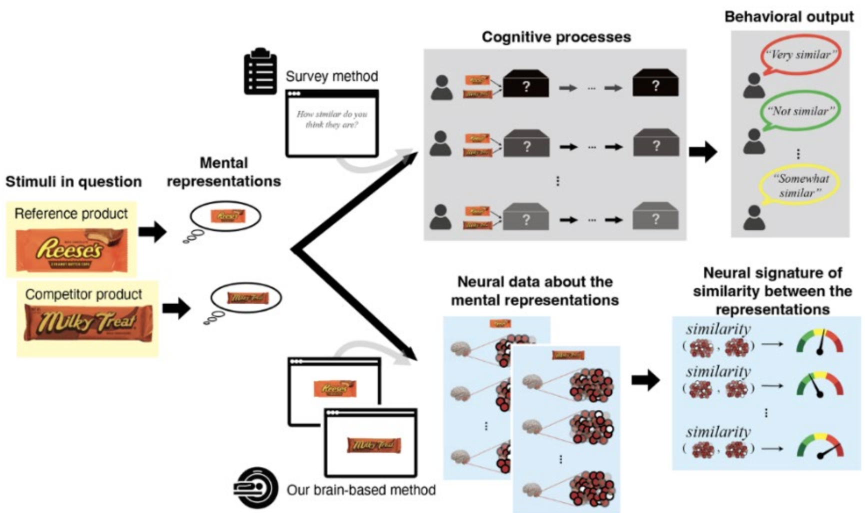

Figure III. Our brain-based measure (bottom) bypasses the complex cognitive processes involved in responding to a survey (top).[51]

[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.