

The Trademark Reporter®



The Law Journal of the International Trademark Association

A New Addition to the Trademark Litigator's Tool Kit: A Neuroscientific Index of Mark Similarity

Mark Bartholomew, Zhihao Zhang, Ming Hsu, Andrew Kayser, and Femke van Horen

Reconstructing the Trademark Registry of Mandate Palestine and What Historical Data Can Reveal

Michael Birnhack

Commentary: Incorporating Uncertainty in Trademark Surveys: Do Respondents Really Know What They Are Talking About?

Barton Beebe, Roy Germano, Christopher Jon Sprigman, and Joel H. Steckel

INTERNATIONAL TRADEMARK ASSOCIATION

675 Third Avenue, New York, NY 10017-5704

Telephone: +1 (212) 642-1700

email: wknox@inta.org

OFFICERS OF THE ASSOCIATION

JOMARIE FREDERICKS.....*President*
DANA NORTHCOTT.....*President-Elect*
ELISABETH BRADLEY.....*Vice President*
DEBORAH HAMPTON.....*Vice President*
DANIEL ZOHNY.....*Treasurer*
SERGIO BARRAGAN.....*Secretary*
ANTHONY DREYER.....*Counsel*
ETIENNE SANZ DE ACEDO.....*Chief Executive Officer*

The Trademark Reporter Committee

EDITORIAL BOARD

EDITOR-IN-CHIEF, CHAIR

RAFFI V. ZEROUNIAN

STAFF EDITOR-IN-CHIEF

WILLARD KNOX

VICE CHAIR

PAMELA CHESTEK

Senior Editors

ANDREW GRAY
KAREN ELBURG
LESLEY GROSSBERG

VERENA VON BOMHARD
DÉSIRÉE FIELDS
BARRY HORWITZ

Director of Legal Resources

LIZ HANELLIN

Senior Staff Editor

BEVERLY HARRIS

Senior Staff Editor

ELIZABETH VENTURO

Editor-in-Chief, Practice Guides and Legal Resources

ROSEMARY DESTEFANO

Composer

BARBARA MOTTER

Editors

TARA ALLSTUN
ABDURRAHIM AYAZ
SARA BAUER
DANIEL BERESKIN
WILLIAM BORCHARD
JEANNETTE CARMADILLA
JOSEPH CARRAFIELLO
TED DAVIS
MICHAEL DENNISTON
MOHAN DEWAN
TUE DO
MATTHEW EZELL
CHRISTINE FELLER
MATHILDE FLORENSON
GABRIELE FOUGNER
ALEX GARENS
STUART GREEN
STACY GROSSMAN
JORDI GÜELL

WINSLOW HALL
JOHN HEINBOCKEL
GANG HU
BRUCE ISAACSON
MANAVI JAIN
SAUMYA KAPOOR
INGRIDA KARINA-BERZINA
ELISABETH KASZNAR FEKETE
CHIAKI KATO
SUSAN KERI
SABRINA LARSON
SCOTT LEBSON
BARBARA LEITAO-JOHNSON
QIANG MA
SONAL MADAN
ILARIA MAGGIONI
MICHAEL MAIER
JAMES MCALLISTER
J. THOMAS MCCARTHY
CATHERINE MITROS

ERIC MORAN
SHANA OLSON
RICCARDO PEROTTI
LUIS HENRIQUE PORANGABA
UDAYVIR RANA
BRANDON RESS
SUSAN RUSSELL
MARTIN SCHMIDT
TOM SCOURFIELD
GIULIO ENRICO SIRONI
SCOTT SLAVICK
MEGAN SNEED
RANDY SPRINGER
DONALD THOMPSON
MARTIN VIEFHUES
JOHN L. WELCH
JASON WHITNEY
ALICJA ZALEWSKA
NANCY ZHANG

Advisory Board

MILES ALEXANDER
WILLIAM BORCHARD
LANNING BRYER
JESSICA CARDON
SANDRA EDELMAN
ANTHONY FLETCHER
ROBERT KUNSTADT

THEODORE MAX
KATHLEEN MCCARTHY
GLENN MITCHELL
JONATHAN MOSKIN
VINCENT PALLADINO
JOHN PEGRAM

ROBERT RASKOFF
PASQUALE RAZZANO
SUSAN REISS
PIER LUIGI RONCAGLIA
HOWARD SHIRE
JERRE SWANN, SR.
STEVEN WEINBERG

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. 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.

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 Strategies, 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 Strategies LLC, <http://www.jaysurveystrategies.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

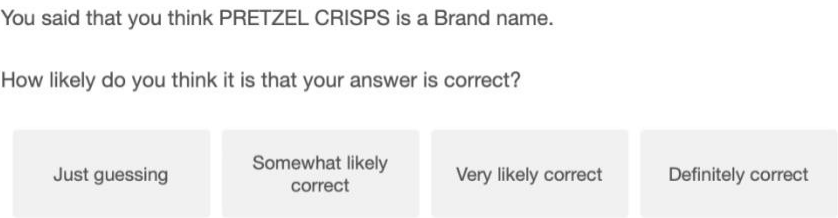
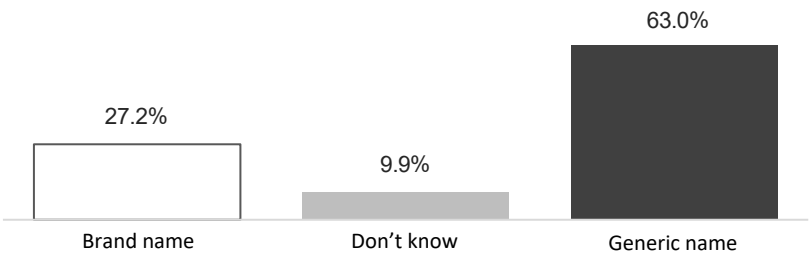


Figure 2 reports the distribution of responses to the first PRETZEL CRISPS question. A clear majority of respondents, 63%, indicated that they believed PRETZEL CRISPS to be a generic term, while just 27% perceived PRETZEL CRISPS as a brand name. Consistent with our experience in other *Teflon* studies, 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.

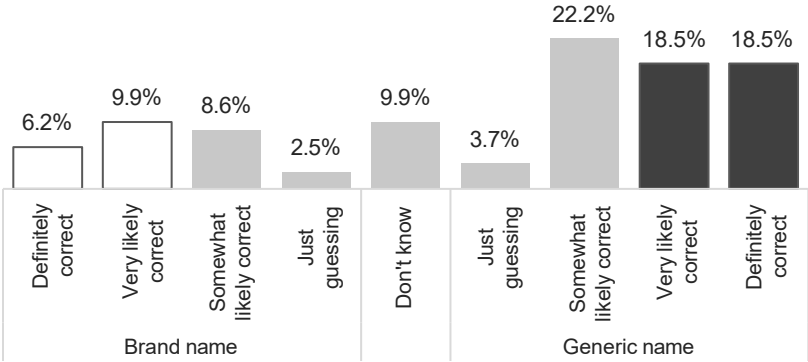
Figure 2
Distribution of Group A Responses to the Standard *Teflon* Question



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

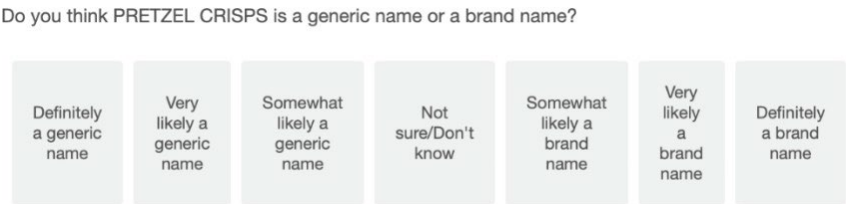
²⁴ See *supra* Part I.B. In *Frito-Lay N. Am., Inc. v. Princeton Vanguard, LLC*, 124 U.S.P.Q.2d 1184 (T.T.A.B. 2017), the Trademark Trial and Appeal Board criticized Dr. Jay’s use in her survey’s initial mini-course of the term WHEAT THINS as an example of a brand, since it is a highly descriptive mark “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.

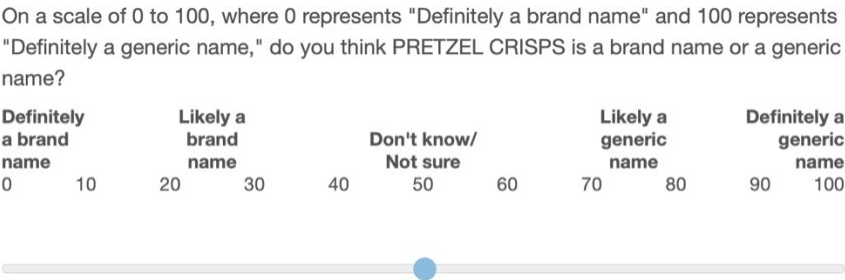
In our *Emory Law Journal* article, you will find discussion of two other ways of assessing respondent uncertainty. In the first, eighty-one respondents were not provided with the “Generic name”/“Brand name”/“Don’t know” answer choices, but rather with a seven-point Likert scale as shown in Figure 4. 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 Burt, an Ohio State University-affiliated psychologist.³⁴ Burt 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. Burt, *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 Bonyne, *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 term⁴⁹ 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

lower⁵²) 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.
