

IN SEARCH OF VALIDITY: A NEW MODEL FOR THE
CONTENT AND PROCEDURAL TREATMENT OF
TRADEMARK INFRINGEMENT SURVEYS

IRINA D. MANTA*

“Objective evidence and certitude are doubtless very fine ideals to play with, but where on this moonlit and dream-visited planet are they found?”—William James

I.	INTRODUCTION.....	1028
II.	SURVEYS AND THE LIKELIHOOD OF CONFUSION.....	1033
	A. <i>The Idea of Likelihood of Confusion</i>	1033
	B. <i>Use of the Survey Method to Establish Likelihood of Confusion</i>	1036
III.	PROBLEMS WITH THE CURRENT TREATMENT OF SURVEY CONTENTS AND FORMATS THROUGH THE EXAMPLE OF THE SECOND CIRCUIT	1041
	A. <i>The Role of Survey Stimuli in the Second Circuit’s Trademark Infringement Jurisprudence</i>	1041
	1. The Creation of the <i>Toys “R” Us</i> Test.....	1042
	2. Survey Stimuli Accepted by the Second Circuit	1043
	a. Surveys Employing the Product Itself or an Image Thereof	1043
	b. The Great Exception: The <i>McBage’s</i> Phone Survey	1045

* Irina D. Manta, Yale Law School, J.D., 2006. Adjunct Professor, University of Arkansas, William H. Bowen School of Law; Law Clerk, Hon. Morris S. Arnold, U.S. Court of Appeals, Eighth Circuit. ©2007 Irina D. Manta.

An earlier draft of this article received the Grand Prize of the Foley & Lardner Intellectual Property Writing Competition. I would like to thank Professor Henry Smith for his guidance, and Professor Douglas Lichtman, Will Baude, Jennifer Peresie, and Matthew Craig for their insights. I would also like to thank Professors Marcus Cole and Todd Zywicki for their comments and the Institute for Humane Studies for its support of this project, with special thanks to Nigel Ashford and the Summer Graduate Research Fellows Class of 2006. Additionally, I owe much to Patterson, Belknap, Webb & Tyler, L.L.P. and in particular to Kim Landsman for his comments and Naomi Schrag for raising my initial interest in this topic.

3.	Survey Stimuli Rejected by the Second Circuit.....	1046
a.	Improper Stimuli and Improper Questions	1046
b.	Lack of a Visual Stimulus	1047
c.	Visual Stimuli Altered from the Original	1049
d.	Inappropriate Visual Control Stimuli.....	1051
e.	Failure to Present Both the Original and Allegedly Infringing Product.....	1052
f.	Compound Problems	1053
g.	Environmental Cues	1054
h.	Leading Brands.....	1055
i.	Business Associations Through Similar Names.....	1056
B.	<i>Summary of Second Circuit Doctrine Regarding Survey Contents and Formats.....</i>	1057
IV.	A NEW MODEL FOR TRADEMARK INFRINGEMENT SURVEYS.	1059
A.	<i>The Procedural Handling of Trademark Surveys.....</i>	1059
B.	<i>Substantive Guidelines for Content-Related Issues.....</i>	1065
1.	The Proper Stimulus	1066
2.	The Proper Control Stimulus	1068
3.	The Proper Questions	1068
4.	The Proper Environment.....	1070
5.	Discussion.....	1070
V.	CONCLUSION	1071

I. INTRODUCTION

Trademark scholars and laymen alike find it suspicious when a company other than Starbucks suddenly starts selling “Charbucks blend” coffee. One’s first intuition is that this is not a coincidence but rather an attempt to capitalize on the name of a pre-existing, immensely popular product and potentially cause losses to the original product through diversion of sales. It hardly comes as a surprise to learn that in this real-life scenario, Starbucks sued the maker of Charbucks products for trademark infringement in the Southern District of New York to prevent further use of the Charbucks name. What is rather remarkable, however, is that despite Starbucks’ hiring of a well-known law firm and attacking a product that even the most casual observer perceives as infringing, Starbucks lost the litigation.¹ What

¹ Starbucks Corp. v. Wolfe’s Borough Coffee, Inc., No. 01 Civ. 5981 (LTS), 2005 WL 3527126 (S.D.N.Y. Dec. 23, 2005).

happened? The district court simply threw out the survey that social science experts hired by Starbucks' legal team had prepared to demonstrate consumer confusion, and without survey evidence it is generally almost impossible to prove trademark infringement.

Considering the case in a vacuum, the outcome of the Starbucks litigation could mean that even good law firms and their experts sometimes make mistakes in survey construction. An overview of trademark jurisprudence in the Second Circuit Court of Appeals, however, indicates a broader problem: the 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.² The lack of certainty with respect to survey evidence risks the erosion of consumer trust in marks and chips away at the signaling function of trademarks, thus preventing marks from properly fulfilling their primary purpose.

A world devoid of trademark protection is difficult to imagine and has in fact barely existed.³ Even marks developed thousands of years ago served the purposes of those present today: trademarks allowed for advertising, answered questions of ownership, and provided quality seals.⁴ In modern language, it is said that trademarks "seek[] to economize on information costs by providing a compact, memorable, and unambiguous identifier of a product or service."⁵ Trademarks thus both serve to attract consumers to a good in the first place and then encourage consumers to return to the product. If an individual liked a particular product, a consistently used mark will give him the assurance that whenever he next encounters a product with the familiar mark, he will enjoy another item that has the same producer and properties as the initial good. Naturally, a producer

² The makers of Starbucks products, of course, believe that the outcome of the case was fair and tell a vastly different story on the issue of trademark infringement. See Black Bear Micro Roastery, *What It's All About?*, Aug. 9, 2006, http://www.blackbearcoffee.com/Starbucks/What's_it_all_about.htm ("How could our 'microscopic' use of the term possibly have any effect on what has been viewed as one of the most spectacular name recognition success stories ever. The preposterous notion that we could possibly slowdown the Starbucks 'freight train' simply didn't occur to us.")

³ ROBERT P. MERGES ET AL., *INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE* 529 (3d ed. 2003) ("Trademarks have existed for almost as long as trade itself."). The first identifiable marks were words or symbols that identified a good's maker and were found on objects as many as four thousand years old. *Id.* (citing WILLIAM H. BROWNE, *A TREATISE ON THE LAW OF TRADEMARKS* 1-14 (1885)). For a more extensive account of the history of trademarks, see generally Thomas D. Drescher, *The Transformation and Evolution of Trademarks—From Signals to Symbols to Myth*, 82 TRADEMARK REP. 301 (1992).

⁴ MERGES ET AL., *supra* note 3, at 529.

⁵ Thomas R. Lee, *Demystifying Dilution*, 84 B.U. L. REV. 859, 887 n.190 (2004) (quoting Richard A. Posner, *When Is Parody Fair Use?*, 21 J. LEGAL STUD. 67, 75 (1992)).

will in turn potentially reap rich financial rewards for distributing popular, high-quality products because customers will return to buy more of the same.⁶ This relationship between producers, consumers, and trademarks can only operate properly, however, if a particular trademark designates only a single product or origin.⁷ Imagine, as an extreme case, a consumer buying a bottle with the familiar Coca Cola label only to find herself drinking a non-caffeinated, strawberry-flavored beverage that was made by a company of a much lower reputation. Indeed, “brand names serve as information ‘chunks.’ They represent core nodes in memory around which other ‘associated’ information is connected and organized. Given only a familiar brand name, a host of relevant and important information can be efficiently called into consciousness.”⁸ When this link between brand names and memory is disrupted because, for instance, another product falsely evokes the memories one associates with the original good, consumer confusion and disappointment ensue.⁹

To avoid situations of consumer confusion and the subsequent potential consumer dissatisfaction and losses to trademark owners, Congress passed legislation against trademark infringement.¹⁰ Most famously, Congress enacted the Lanham Act¹¹ in 1946 to prevent trademark infringement and false advertising. Since the inception of the Act, the extent of intellectual property protection offered to trademark owners has greatly increased.¹² One of the most important additions to the

⁶ See, e.g., Joshua Kennon, *Starbucks Franchise Value—A Case Study of an Excellent Business*, ABOUT.COM, Jan. 17, 2005, <http://beginnersinvest.about.com/cs/newinvestors/a/021103a.htm> (attributing the fact that “from 1997 to 2001, [Starbucks’s] revenues increased from \$975,389,000 to over \$2,648,980,000” to the company’s “powerful brand name.”).

⁷ William M. Landes & Richard A. Posner, *Trademark Law: An Economic Perspective*, 30 J.L. & ECON. 265, 269 (1987) (“To perform its economizing function a trademark or brand name . . . must not be duplicated.”).

⁸ Jacob Jacoby, *The Psychological Foundations of Trademark Law: Secondary Meaning, Genericism, Fame, Confusion and Dilution*, 91 TRADEMARK REP. 1013, 1024-25 (2001).

⁹ But see Stacey L. Dogan, *An Exclusive Right to Evoke*, 44 B.C. L. REV. 291 (2003) (arguing against the dangers of equating evocation with infringement).

¹⁰ See, e.g., 1 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 2:33 (4th ed. 1996) (“Almost all trademark disputes are between firms that use conflicting marks; the consumer is not a party to the litigation. But it is the consumer’s state of mind that largely controls the result”) (footnote omitted). Professor McCarthy explains that trademark infringement both “inhibits competition” and “deprives consumers of their ability to distinguish among the goods of competing manufacturers.” *Id.* (citing *Inwood Labs., Inc. v. Ives Labs., Inc.*, 456 U.S. 844, 855 n.14 (1982)).

¹¹ 15 U.S.C. §§ 1051-1052 (2006).

¹² WILLIAM M. LANDES & RICHARD A. POSNER, *THE POLITICAL ECONOMY OF INTELLECTUAL PROPERTY LAW* 2 (2004), available at <http://aei-brookings.org/admin/authorpdfs/page.php?id=985> (using the number of words in the Lanham Act and other intellectual property statutes as a “crude measure of the expansion in intellectual property rights . . . since most of these statutes expand such rights or create

initial text of the Lanham Act was the 1995 passage of the Federal Trademark Dilution Act (FTDA).¹³ The FTDA, which did not require a showing of likelihood of consumer confusion, was designed to protect against uses of famous marks that would lead to the tarnishment of marks' image as well as the blurring of marks' distinctiveness.¹⁴

Courts have allowed the introduction of several types of evidence to demonstrate that consumer confusion has occurred in violation of the Lanham Act, the specific test consisting of whether there is a "likelihood of confusion."¹⁵ Likelihood of confusion can be shown by demonstrating that the defendant intentionally imitated or copied the plaintiff's mark,¹⁶ that there was actual consumer confusion in the marketplace,¹⁷ and, most crucially, through consumer surveys.¹⁸ The importance of consumer surveys in the context of trademark litigation cannot be overemphasized:¹⁹

new ones rather than reduce existing rights," and noting that the Lanham Act "had 10,640 words in 1946, 13,345 in 1987, a jump to 20,136 in 1988 with the passage of the Trademark Revision Act, and 24,750 in 2000."); see also William M. Landes, *An Empirical Analysis of Intellectual Property*, 41 Hous. L. Rev. 749, 750 n.1 and accompanying text (2004) (citing LANDES & POSNER, *supra* at 4 and noting that there has been little increase in intellectual property litigation over the past twenty-five years but great growth in the "output and legal protection of intellectual property").

¹³ 15 U.S.C. § 1125(c).

¹⁴ For a short description of the FTDA, see *Trademark Dilution*, BITLAW, Jan. 17, 2005, <http://www.bitlaw.com/trademark/dilution.html>. Prior to the enactment of the FTDA, trademarks were often protected against dilution through a number of state statutes. Many states modeled their anti-dilution statutes after the Model State Trademark Bill which was developed by the International Trademark Association, available at <http://www.inta.org/policy/mstb.html> (last visited Jan. 17, 2005). For critiques of current trademark dilution doctrine, see Jonathan E. Moskin, *Dilution or Delusion: The Rational Limits of Trademark Protection*, 83 TRADEMARK REP. 122 (1993).

¹⁵ 3 MCCARTHY, *supra* note 10, § 23:12. McCarthy also states that courts have consistently held that, "[t]o prove liability, the plaintiff is not required to prove any instances of actual confusion." *Id.* (footnote omitted).

¹⁶ 2 *id.* § 15:38 ("Evidence that defendant knowingly imitated or copied plaintiff's symbol has long been regarded as probative of likely customer confusion.") (footnote omitted). But see Timothy R.M. Bryant, *Trademark Infringement: The Irrelevance of Evidence of Copying to Secondary Meaning*, 83 NW. U. L. REV. 473, 511 (1989) (concluding that evidence of copying should not be admitted in Lanham Act cases) (cited in 2 MCCARTHY, *supra* note 10, § 15:38 n.10).

¹⁷ 3 MCCARTHY, *supra* note 10, § 23:13 ("Convincing evidence of significant actual confusion occurring under actual marketplace conditions is the best evidence of a likelihood of confusion."). See generally Michael J. Allen, *The Role of Actual Confusion Evidence in Federal Trademark Infringement Litigation*, 16 CAMPBELL L. REV. 19 (1994) (cited in 3 MCCARTHY, *supra* note 10, § 23:13 n.1).

¹⁸ 5 MCCARTHY, *supra* note 10, § 32:195 (calling surveys "the most direct method of demonstrating secondary meaning and likelihood of confusion") (citing *Charles Jaquin Et Cie, Inc. v. Destileria Serralles, Inc.*, 921 F.2d 467, 476 (3d Cir. 1990)). A number of courts have chosen to "treat[] market survey evidence as evidence of actual confusion even though it is limited to responses to the artificially created survey stimuli." Allen, *supra* note 17, at 27-28.

¹⁹ Jack P. Lipton, *Trademark Litigation: A New Look at the Use of Social Science Evidence*, 29 ARIZ. L. REV. 639 (1987). The author states:

In the trademark area, social science surveys, generally pertaining to the state of mind of prospective purchasers, have been offered as evidence of the existence of secondary meaning or consumer confusion. Surveys have also been used

intellectual property protection for trademarks focuses on potentially false impressions created in market actors, and surveys constitute the main tool to measure “the mental state of some segment of the consuming public.”²⁰ As is often the case with social science, however, many questions have arisen as to the proper construction of surveys used in trademark litigation.²¹ This article focuses on some of the issues that have provoked controversy in a significant number of cases that employed trademark-related surveys, such as what the stimulus in consumer surveys should consist of, and what test courts should articulate for whether the proper stimulus has been used. Stimuli are commonly understood as the materials that survey subjects face that are supposed to represent the potentially infringing product or service. This article argues that courts’ treatment of acceptable survey stimuli has been plagued by inconsistencies and that courts need to more clearly elucidate *ex ante* rules of what will constitute proper stimuli to demonstrate Lanham Act violations. This article also discusses the standards for survey components directly related to survey stimuli, namely, control stimuli, environmental cues, and the formulation of proper questions, and strives to provide a consistent model for the contents and formats of surveys.

Part II of this article will expand on the use of surveys in trademark-related cases and explain the general principles of correct survey construction. Part III will scrutinize the theoretical underpinnings of survey contents as well as how courts treat surveys, with a particular focus on cases in the Second Circuit Court of Appeals. An examination of Second Circuit cases is especially appropriate for this type of analysis because the Second Circuit handled a significant number and perhaps the most trademark-related cases of any Circuit.²² Part III will also attempt

specifically to assess whether a brand name has become generic. Finally, parties have attempted to introduce survey evidence in support of or in opposition to an application for federal trademark registration. It has been stated that the evaluation of survey evidence is one of the most controversial and difficult problems confronting the Trademark Trial and Appeal Board and other judicial entities dealing with trademark cases. Nevertheless, surveys, as a means of assessing consumers’ state of mind, have played an essential role in aiding courts to make factual findings under the appropriate legal test.

Id. at 641-42 (footnotes omitted).

²⁰ *Id.* at 642.

²¹ See, e.g., Susan J. Becker, *Public Opinion Polls and Surveys as Evidence: Suggestions for Resolving Confusing and Conflicting Standards Governing Weight and Admissibility*, 70 OR. L. REV. 463 (1991).

²² A search of Westlaw’s court of appeals database (CTA) and district court database (DCT) with the synopsis search “sy(trademark!)” found many more trademark-related cases in the Second Circuit than in any other circuit, with the exception of the Federal Circuit. There were 1,332 cases in the Second Circuit as opposed to the next largest number, 542 cases, found in the Ninth Circuit, between 143 and 485 cases in the other circuits, and 1,687 in the Federal Circuit. While this is a very rough estimate and contains

to demonstrate the obstacles that courts in the Second Circuit have faced when determining the validity of survey contents and critique tensions between different decisions of the court. Part IV proposes a model for rules regarding surveys that courts should implement that both conform to the Rule of Law and serve as clearer guidelines than has been exemplified thus far in the jurisprudence. This Part also discusses alternative approaches to resolving survey issues, such as pre-clearance of surveys through the Patent & Trademark Office (PTO) and the creation of congressional guidelines for surveys. Part V concludes by emphasizing the need to apply the suggested changes to trademark jurisprudence.

II. SURVEYS AND THE LIKELIHOOD OF CONFUSION

A. *The Idea of Likelihood of Confusion*

The original Lanham Act of 1946 stated that using a mark “likely to cause confusion or mistake or to deceive purchasers as to the source or origin of such goods or services”²³ constitutes trademark infringement. While the legislative history of the Lanham Act does not directly address Congress’ choice of those words, it does clearly state the overall purpose of the law: to “protect the public so it may be confident that, in purchasing a product bearing a particular trade-mark which it favorably knows,

several errors of measurement, the significant disparity in numbers does tend to suggest that the Second Circuit probably decides the greatest numbers of trademark-related cases in the country. There are several reasons to employ the Second Circuit for analysis rather than the Federal Circuit. This includes the fact that the database search here displays particularly unreliable results for the Federal Circuit as illustrated from the count of the numbers of search results that just involved the “Patent and Trademark Office” generally.

A number of reasons have led to the large number of trademark suits in the Second Circuit. For one, many industries are located in New York, such as the media, entertainment, fashion, and toys and games businesses, all of which generate trademark disputes. In terms of industries, New York City and Los Angeles present the most logical locations for such suits. Many trademark lawyers prefer to initiate actions in New York City rather than Los Angeles because one can obtain preliminary injunctions in the Southern District of New York within a very short time, and because the Second Circuit is considered more reasonable and consistent than the Ninth Circuit, in which Los Angeles is located.

Another motivation for exploring Second Circuit precedent is that several of the most significant appellate court cases regarding trademark infringement that inspired the jurisprudence of courts across the country originated in the Second Circuit. *See, e.g.*, *Polaroid Corp. v. Polarad Elecs. Corp.*, 287 F.2d 492 (2d Cir. 1961) (providing a test to determine the presence of trademark infringement); *Toys “R” Us, Inc. v. Canarsie Kiddie Shop, Inc.*, 559 F. Supp. 1189 (E.D.N.Y. 1983) (defining a test for the trustworthiness of trademark infringement surveys); *Zippo Mfg. Co. v. Rogers Imps., Inc.*, 216 F. Supp. 670 (S.D.N.Y. 1963) (establishing that the hearsay rule does not bar the admission of consumer surveys).

²³ Lanham Act, ch. 540, § 32(1)(a), 60 Stat. 427, 437 (1946) (current version at 15 U.S.C. § 1114(1) (2006)).

it will get the product which it asks for and wants to get”²⁴ and to ensure the trademark owner is “protected in his investment from its misappropriation by pirates and cheats.”²⁵ Congress also emphasized its desire to “eliminate judicial obscurity . . . and [make] relief against infringement prompt and effective.”²⁶ Any analysis or critique of current trademark infringement jurisprudence must consider these expressions of legislative intent.

Courts have had to make the statutory concept of likelihood of confusion operational in a judicial setting. They have had to define both what it means for a mark to be likely to cause confusion and how trademark owners can prove the existence of this confusion. One trademark expert elaborates on these points:

Likelihood of confusion concerns the probable or actual actions and reactions of people, making it fundamental to understand how they make their decisions to buy or not to buy. A realistic evaluation of consumer confusion must attempt to recreate the conditions in which buying decisions are made, and the court should try to determine not what it would do, but what a reasonable purchaser in market conditions would do.²⁷

Two of the most important aspects of substantiating a claim based on likelihood of confusion can be summarized as requiring 1) a demonstration of a particular subjective state of mind on the part of consumers and 2) that the potential confusion is one that would take place in the actual marketplace.²⁸ Trial courts must show particular care when deciding how to evaluate evidence surrounding likelihood of confusion because their rulings on these issues will constitute factual findings subject to the “clearly erroneous” standard if the case reaches an appellate court.²⁹ Such a high barrier means that if a trial court rules incorrectly on a matter of likelihood of confusion, the wronged party may never be

²⁴ H.R. REP. NO. 79-219, at 2 (1945); *accord* S. REP. NO. 79-1333, at 3 (1946).

²⁵ H.R. REP. NO. 79-219, *supra* note 24, at 2.

²⁶ *Id.*

²⁷ RICHARD L. KIRKPATRICK, LIKELIHOOD OF CONFUSION IN TRADEMARK LAW § 2.2 (2006) (footnotes omitted).

²⁸ Kirkpatrick emphasizes this point further:

The court is not concerned with mere theoretical possibilities of confusion, deception, or mistake or with de minimis situations but with the practicalities of the commercial world, with which the trademark laws deal. The realities of the marketplace control because that is where confusion of prospective purchasers would or would not occur. The public’s perspective controls, not the court’s.

Id. (internal quotation marks and footnotes omitted).

²⁹ *See* 3 MCCARTHY, *supra* note 10, § 23:69 (explaining that Rule 52(a) of the Federal Rules of Civil Procedure imposes the “clearly erroneous” standard on courts of appeals) (citing *Inwood Labs., Inc. v. Ives Labs., Inc.*, 456 U.S. 844, 857-58 (1982) (“An appellate court cannot substitute its interpretation of the evidence for that of the trial court simply because the reviewing court might give the facts another construction, resolve the ambiguities differently, and find a more sinister cast to actions which the District Court apparently deemed innocent.”)).

able to obtain a remedy for the court's mistake.

Over the years, Congress cast a relatively wide net to capture many different kinds of confusion. One source explains that "Congress 'evinced a clear purpose to outlaw the use of trademarks which are likely to cause confusion, mistake, or deception of any kind, not merely of purchasers nor simply as to source of origin.'"³⁰ The types of confusion that give rise to a cause of action include "source confusion,"³¹ "sponsorship confusion,"³² "reverse confusion,"³³ "subliminal and associational confusion,"³⁴ and "confusion of marks and of products."³⁵ Courts have created a number of multi-factor tests to determine whether there is likelihood of confusion. One of the most famous standards is the one developed by the Second Circuit in *Polaroid Corp. v. Polarad Electronics Corp.*,³⁶ where the court established an eight-factor test for infringement:

Where the products are different, the prior owner's chance of success is a function of many variables: [1] the strength of his mark, [2] the degree of similarity between the two marks, [3] the proximity of the products, [4] the likelihood that the prior owner will bridge the gap, [5] actual confusion, and [6] the reciprocal of defendant's good faith in adopting its own mark, [7] the quality of defendant's product, and [8] the sophistication of the buyers.³⁷

While the other circuits use slightly modified factors, the tests tend to be much more similar than different,³⁸ and all circuits

³⁰ KIRKPATRICK, *supra* note 27, at 129 (quoting *Syntex Labs., Inc. v. Norwich Pharmacal Co.*, 437 F.2d 566, 568 (2d Cir. 1971)).

³¹ *Id.* § 1.4.A (explaining that a source "is the entity which controls the nature and quality of the product") (footnote omitted).

³² *Id.* § 1.4.B (labeling "confusion about sponsorship, approval or certification by one source of the goods or services of another, or about an affiliation, connection or association between one source and another" as sponsorship confusion) (footnote omitted).

³³ *Id.* § 1.4.C (explaining that "reverse confusion is the misimpression that the newcomer is the source or sponsor of the prior user's goods") (footnote omitted).

³⁴ Kirkpatrick states:

The consumer may subliminally or subconsciously identify the properties and reputation of one product with those of another, although he can identify the particular manufacturer of each. Even if consumers do not consciously assume that the defendant's product is somehow affiliated with plaintiff's product, there is the likelihood that consumers will be attracted to defendant's product on the strength of the goodwill and positive image established by plaintiff. Thus, the newcomer may gain a foothold in the first user's market by exploiting subliminal associations with the prior mark.

Id. § 1.4.D (internal quotation marks and footnotes omitted).

³⁵ *Id.* § 1.4.E (stating that "confusion may relate to the public's knowledge that the trademark, which is 'the triggering mechanism' for the sale of the product, originates with the plaintiff") (footnote omitted).

³⁶ *Polaroid Corp. v. Polarad Elecs. Corp.*, 287 F.2d 492 (2d Cir. 1961).

³⁷ *Id.* at 495.

³⁸ For an excellent summary of the tests adopted by each circuit, see KIRKPATRICK, *supra* note 27, § 2.4. The different tests generally include the following common factors:

allow the use of trademark survey evidence. Although this article focuses on the jurisprudence of the Second Circuit Court of Appeals, the proper implementation of the federal statutory framework of the Lanham Act inherently poses similar problems for every circuit.³⁹ Thus, the solutions proposed in this article are targeted towards adoption across circuits.

B. *Use of the Survey Method to Establish Likelihood of Confusion*

As previously mentioned, it is critical to examine the mindset of actual or potential customers to determine their level of confusion about a possibly infringing product. Unfortunately, a judge cannot just stop people and directly ask them whether they would feel confused about a good or not. Most direct questions would either bias or confuse the customer (“are you likely to be confused as to who made this bottle of soda?”) and would not reach individuals’ subconscious cognitive processes that often affect their decisions. Thus, courts often need to rely on consumer survey evidence as an indirect measure of consumer confusion. In fact, McCarthy notes that “[a]s the use of surveys has become more common, judges have to come to expect that a survey will be introduced to aid the court in determining customers’ state of mind” and the lack of a survey is sometimes taken as a sign “that a litigant is less than deadly serious about its case.”⁴⁰

Consumer surveys are conducted in a number of forms and media. As a general matter, one or both of the litigants hire social science experts to conduct surveys that provide evidence for or against the existence of confusion on the part of potential product or service users. To do so, experts present consumers with a representation of the allegedly infringing product and ask them a number of questions that indirectly measure the level of confusion (e.g., who the consumers think made the product at issue, what

“the strength of the senior mark; the similarity of the marks; the relatedness of the products and marketing channels; the consumers and their degree of care; actual confusion; and the junior user’s intent.” *Id.* For a summary of the factors used in different circuits in tabular form, see Edward George Epstein, *Surveys: Growing Admissibility but Narrow Utilization*, 83 TRADEMARK REP. 863, 871 (1993).

³⁹ For further evidence showing that different circuits have encountered similar problems when deciding the admissibility and weight of various surveys, see Daniel A. Klein, *Admissibility and Weight of Consumer Survey in Litigation Under Trademark Opposition, Trademark Infringement, and False Designation of Origin Provisions of Lanham Act (15 U.S.C.A. §§ 1063, 1114, and 1125)*, 98 A.L.R. FED. 20 (2005).

⁴⁰ 5 MCCARTHY, *supra* note 10, § 32:195; *see also* Citigroup Inc. v. City Holding Co., 99 Civ. 10115 (RWS), 2003 U.S. Dist. LEXIS 1845, at *72 (S.D.N.Y. Feb. 10, 2003) (quoting Merriam-Webster, Inc. v. Random House, 35 F.3d 65, 72 (2d Cir. 1994) (“The lack of survey evidence counts against finding actual confusion.”)); E.S. Originals Inc. v. Stride Rite Corp., 656 F. Supp. 484, 490 (S.D.N.Y. 1987) (“[T]hat [plaintiff] did not undertake a consumer survey . . . strongly suggests that a likelihood of confusion cannot be shown.”).

other products originate from the same maker, etc.). Experts record consumers' answers and perform statistical analyses on the survey results. Usually, surveys also contain representations of control products such that experts can compare the amount of confusion that occurs with the allegedly infringing product against the baseline of how much confusion even an entirely neutral-looking product would cause. While surveys differ in their execution, number and exact formulation of questions, and product representations, they usually share basic characteristics.⁴¹

In early trademark cases, judges tended to reject the survey method, partly based on hearsay concerns, or asked for live witnesses in addition to survey results.⁴² In the end, however, courts "recognized that survey evidence was not proof of the truth of the survey respondents' assertions but instead was offered simply to show their state of mind when confronted with the survey stimulus."⁴³ Once this barrier had been cleared, courts generally accepted surveys as evidence and determined that questions as to the credibility that a survey should be accorded needed to go to the weight rather than the admissibility of the survey.⁴⁴ Thus, it sometimes became the jury's task to decide whether a survey was trustworthy.⁴⁵ This standard radically changed with the decision in *Daubert v. Merrell Dow Pharmaceuticals*,⁴⁶ which recognized a "gatekeeping role for the judge"⁴⁷ in deciding whether to admit expert evidence.⁴⁸ The

⁴¹ For an explanation of some of the main survey formats that experts use, see Itamar Simonson, *The Effect of Survey Method on Likelihood of Confusion Estimates: Conceptual Analysis and Empirical Test*, 83 TRADEMARK REP. 364 (1993).

⁴² Lipton, *supra* note 19, at 644-47. "The hearsay rule may still apply if the people who actually conducted the surveys are not available for cross-examination in either trial or deposition." *Id.* at 646 (citations omitted).

⁴³ G. Kip Edwards, *Lanham Act Surveys After Daubert: Lessons Learned So Far*, 677 PLI/PAT 421, 423 (2001) (citing 5 MCCARTHY, *supra* note 10, §§ 32:270-32:271). Edwards adds: "The adoption of the Federal Rules of Evidence in 1975, and specifically Rule 703, put to rest any hearsay objection to survey evidence, at least when the evidence was offered to prove state of mind." *Id.*

⁴⁴ *Id.* at 424.

⁴⁵ Commenting on procedure in trademark infringement and unfair competition litigation, McCarthy states: "In the federal courts, the right to trial by jury for civil legal actions of over \$20 is guaranteed by the Seventh Amendment to the U.S. Constitution. In the state courts, the right to trial by jury in legal actions is guaranteed by the state constitution of most states." 5 MCCARTHY, *supra* note 10, § 32:122 (citations omitted). While plaintiffs do not always request jury trials, the number of jury trials in the trademark infringement context greatly increased in the 1970s. *Id.* "The publicity given to the multi-million dollar jury verdict in the BIGFOOT case may have had something to do with the increasing popularity of jury trials in trademark cases." *Id.* § 32:122 n.3 (quoting *Big O Tire Dealers, Inc. v. Goodyear Tire & Rubber Co.*, 408 F. Supp. 1219 (D. Colo. 1976)).

⁴⁶ *Daubert v. Merrell Dow Pharm. Inc.*, 509 U.S. 579 (1993).

⁴⁷ *Id.* at 597.

⁴⁸ For a more extended discussion of *Daubert's* gate-keeping function, see Richard J. Leighton, *Using Daubert-Kumho Gatekeeping to Admit and Exclude Surveys in Lanham Act Advertising and Trademark Cases*, 92 TRADEMARK REP. 743, 747-53 (2002).

Daubert Court explained that “the trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.”⁴⁹ The Supreme Court later extended *Daubert* from scientific expert testimony to “all expert testimony.”⁵⁰ On appeal, so-called “*Daubert* motions” that attempt to exclude unreliable evidence are subject to the “abuse of discretion” standard.⁵¹ Endowing courts with this gatekeeping function has increased the pressure on judges to objectively determine what constitutes acceptable evidence. Mistakenly throwing out a survey in a trademark case could cause irreparable damage to a plaintiff given the central nature of the survey in proving likelihood of confusion. Denying a *Daubert* motion, however, sends a signal that the judge lends at least some credibility to the survey and that a jury would do well to use the survey results in their decision. Thus, judges cannot simply use an overwhelmingly strong presumption of trustworthiness for surveys—the fact that a decision on the *Daubert* motion could decide the whole case simply raises the stakes too high. It is generally understood that “no survey is perfect” and that some problems “in questions and methodology should only affect the weight accorded survey results”⁵² rather than admissibility under *Daubert*. At the same time, it is troublesome if many surveys are admitted that are later accorded little to no weight by the jury or judge.⁵³

⁴⁹ *Daubert*, 509 U.S. at 589.

⁵⁰ Edwards, *supra* note 43, at 426 n.2 (citing *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999)).

⁵¹ *Id.* Judge Kozinski of the Ninth Circuit Court of Appeals quipped that the “abuse of discretion standard . . . basically asks, ‘Would the trial judge have to have been drunk or crazy to make this decision?’ If we think a sober person could come to that conclusion, then we have to let it stand.” (quoting Alex Kozinski, *Expert Testimony After Daubert*, 192 J. ACCT. ONLINE 59, 60 (2001), available at <http://www.aicpa.org/pubs/jofa/jul2001/kozinski.htm>).

⁵² 5 MCCARTHY, *supra* note 10, § 32:170 (quoting *Selchow & Righter Co. v. Decipher, Inc.*, 598 F. Supp. 1489, 1502 (E.D. Va. 1984)).

⁵³ One commentator states:

In Lanham Act cases, expert testimony in the form of surveys and related opinion is regularly proffered to prove how advertisements or trademarks influence perceptions. Since at least 1960, federal guidelines have urged trial judges to conduct a pretrial inquiry into the reliability of proffered expert survey evidence. Nonetheless, many courts and influential reference works have employed, without substantive explanation, an admit-it-now-worry-about-it-later presumption for surveys that contain alleged methodological or technical faults. Many courts simply admit allegedly faulty surveys into evidence for later determination of their weight, if any at all. Reviewers have determined that surveys in Lanham Act false advertising cases “appear almost never to be rejected,” yet 23 percent of them were accorded little or no weight by the factfinder (the jury or the judge in a non-jury proceeding).

Leighton, *supra* note 48, at 743-44 (footnotes omitted) (citing Jacob Jacoby et al., *Survey Evidence in Deceptive Advertising Cases Under the Lanham Act: An Historical View of Comments from the Bench*, 84 TRADEMARK REP. 541, 549-50 (1994)); see also Kenneth A. Plevan, *Daubert’s Impact on Survey Experts in Lanham Act Litigation*, 95 TRADEMARK REP. 596, 605-06 (2005) (examining forty-four cases that contained Lanham Act surveys between 1997 and

Daubert and its progeny had a profound impact on trademark litigation. Attorneys began to file and win motions to exclude their opponents' survey evidence prior to trial.⁵⁴ Courts, however, exercised their gatekeeping function by making what one expert called "decisions that border on the schizophrenic."⁵⁵ The reasons for this predicament are manifold, but at its root is the complexity of the task itself, which Judge Kozinski on the Ninth Circuit pointed out on remand in the *Daubert* case:

Our responsibility, then, unless we badly misread the Supreme Court opinion, is to resolve disputes among respected, well-credentialed scientists about matters squarely within their expertise, in areas where there is no scientific consensus as to what is and what is not "good science," and occasionally to reject such expert testimony because it was not "derived by the scientific method."⁵⁶

After some lower courts attempted to limit the effect of *Daubert* to only certain classes of "scientific" expert evidence,⁵⁷ the Supreme Court clarified in *Kumho Tire Co. v. Carmichael*⁵⁸ that the framework should apply to all types of expert evidence. While doing so already proved difficult for purely scientific questions, the inherently vague nature of evidence of a *social* scientific nature certainly did not simplify matters. Social science tends to allow more room for expert bias and does not always easily lend itself to falsification.⁵⁹ The importance of resolving questions surrounding

2004 and finding that courts excluded or gave no weight to the "not-insignificant number" of fourteen of them).

⁵⁴ Leighton, *supra* note 48, at 745-46.

⁵⁵ *Id.* at 746 (footnote omitted).

⁵⁶ *Daubert v. Merrell Dow Pharms.*, 43 F.3d 1311, 1316 (9th Cir. 1995). Judge Kozinski calls this new framework a "Brave New World." *Id.* Others have also noted similar difficulties on the part of lower courts to implement *Daubert*. See, e.g., David L. Faigman, *The Law's Scientific Revolution: Reflections and Ruminations on the Law's Use of Experts in Year Seven of the Revolution*, 57 WASH. & LEE L. REV. 661, 664-69 (2000). But see Laurens Walker & John Monahan, *Daubert and the Reference Manual: An Essay on the Future of Science in Law*, 82 VA. L. REV. 837, 857 (1996) (noting the positive consequences of the *Daubert* decision, such as how it will improve the ways in which "courts can learn about the real world").

⁵⁷ Faigman, *supra* note 56, at 665.

⁵⁸ *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999). After the *Kumho Tire* decision, the scholarly community often used the term "*Daubert-Kumho* framework" to refer to *Daubert* and its progeny. An amendment to the Federal Rules of Evidence later codified the requirements of this framework. FED. R. EVID. 702 (requiring testimony experts to be "based upon sufficient facts or data" and "the product of reliable principles and methods," and requiring the expert witness to have "applied the principles and methods reliably to the facts of the case").

⁵⁹ One scholar states that the psychology field "is somewhat of a hybrid While most people would admit that psychological research is more 'scientific' than other social sciences such as history or sociology, most would also contend that it is nowhere as 'scientific' as the physical sciences, such as biology, chemistry, and physics." David McCord, *Syndromes, Profiles and Other Mental Exotica: A New Approach to the Admissibility of Nontraditional Psychological Evidence in Criminal Cases*, 66 OR. L. REV. 19, 29 (1987).

admission of expert social science has perhaps nowhere been as dramatic as in the context of criminal law, where decisions on the reliability of data concerning eyewitness identification, battered woman syndrome, repressed memories, and many other areas can have a significant impact on questions of life and death, or at least of freedom and incarceration.⁶⁰ While this article does not seek to perform an in-depth exploration of the dilemmas that *Daubert* poses outside the trademark context, it is interesting to note that courts have encountered many of the same problems exercising their gate-keeping function in these other areas. As one expert summarizes the treatment of forensic identification science under *Daubert*:

[T]his area of law is riddled with contradiction, confusion and chaos. Here is an area of law typified by uninformed acceptance or rejection (usually the former) of empirical claims that courts rarely made any effort to try to understand; by judicial opinions that failed to explain the courts' reasoning about the asserted expertise; by evasion or manipulation of the governing legal tests, or the use of no legal test at all⁶¹

It does not appear that courts have found a satisfying response to the challenges that *Daubert* poses to social science outside trademark law, either. Any proposals to modify judicial treatment of trademark infringement surveys will accordingly benefit from fairly little guidance from other legal fields that *Daubert* affects and will have to stand on their own.

A body of scholarship has begun to develop that seeks to guide courts in making objective decisions about the different elements of trademark surveys, both when deciding the admissibility and subsequently when making a decision on the weight of the evidence. Trademark survey expert Michael Rappoport wrote an extensive article on the criteria courts should use to evaluate the validity of surveys.⁶² He warns that surveys are by necessity "imperfect," mainly because people never act exactly the same when surveyed as in real life, the resources available to conduct surveys are limited, and often surveyed individuals just do not know the answer to the crucial questions in a case.⁶³ Nonetheless, many aspects of surveys can be substantially improved if their conductors keep certain guiding principles in

⁶⁰ Some of these fields of social science expertise are mentioned in Faigman, *supra* note 56, at 674.

⁶¹ Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069, 1128 (1998).

⁶² Michael Rappoport, *Litigation Surveys—Social "Science" As Evidence*, 92 TRADEMARK REP. 957 (2002).

⁶³ *Id.* at 960.

mind. First, a surveyor must ask himself what the “universe” of a study should be—who are the relevant people to interview and how should one properly obtain a representative sample from that group?⁶⁴ Broadly speaking, the right type of universe is one that “focuses on potential purchasers of the defendant’s products or services,”⁶⁵ but several other issues that are beyond the scope of this article must be considered as well.⁶⁶ Next, the surveyor must determine the content of the study, which includes choosing an interview setting, defining the questions to ask and their order, and selecting survey stimuli as well as control stimuli.⁶⁷ Some of these decisions overlap in significant ways. For example, conducting a phone-only survey just allows for the use of auditory stimuli while a mail-in survey can only convey visual elements.⁶⁸

The next Part will focus on the survey contents and formats the Second Circuit Court of Appeals has deemed acceptable. The analysis of the related cases will show that courts need to establish clearer rules for what they deem acceptable options and will conclude with a proposal of how they should go about that task. This Part also discusses alternative ideas involving the potential role of the PTO and legislative guidance to resolve these issues.

III. PROBLEMS WITH THE CURRENT TREATMENT OF SURVEY CONTENTS AND FORMATS THROUGH THE EXAMPLE OF THE SECOND CIRCUIT

A. *The Role of Survey Stimuli in the Second Circuit’s Trademark Infringement Jurisprudence*

What constitutes proper survey stimuli has been one of the key issues plaguing the jurisprudence on survey contents. The general understanding has been that survey stimuli must either be 1) the potentially infringing product itself or 2) a representation thereof that resembles the potentially infringing product in all the

⁶⁴ *Id.* at 963-65. Many of the surveys that fail scrutiny do so because of an improperly defined universe. See, e.g., *Winning Ways, Inc. v. Holloway Sportswear, Inc.*, 913 F. Supp. 1454, 1467 (D. Kans. 1996) (calling a survey definition of the consumer market universe “fatally underinclusive”) (cited in Edwards, *supra* note 43, at 429-30). For further examples of surveys that displayed a defective universe, see Sandra Edelman, *How to Construct a Valid Confusion Survey*, 720 P.L.I./PAT 391, 397-405 (2002).

⁶⁵ See Edelman, *supra* note 64, at 405.

⁶⁶ For a summary of factors to consider in survey universe construction, see *id.* For further discussion on the topic of proper universe construction, see Robert C. Bird, *Streamlining Consumer Survey Analysis: An Examination of the Concept of Universe in Consumer Surveys Offered in Intellectual Property Litigation*, 88 TRADEMARK REP. 269 (1998); Shashank Upadhye, *Trademark Surveys: Identifying the Relevant Universe of Confused Consumers*, 8 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 549 (1998).

⁶⁷ See, e.g., Rappeport, *supra* note 62, at 978.

⁶⁸ A discussion of the stimuli-related problem created by interview structures that are not directly in person will follow *infra* Part III.

ways that are relevant to judge the existence of infringement. Surveys generally also contain corresponding control stimuli such that researchers can compare the level of confusion of the potentially infringing product to a basic threshold.⁶⁹ As will become clear below, survey and control stimuli have consisted of such diverse items as actual products, modified products, images of products, and spoken words. Courts have often struggled with deciding what stimuli to allow, which is ultimately due to the complexity of the relationship between human cognition and the number of elements that make up a trademarked product. To analyze how courts have treated survey stimuli, this Part examines the doctrine of the Second Circuit Court of Appeals, which has dealt extensively with the issue.⁷⁰

1. The Creation of the *Toys “R” Us* Test

In determining the validity of surveys used in the trademark context, the Second Circuit has widely adopted the test set out in *Toys “R” Us, Inc. v. Canarsie Kiddie Shop, Inc.*,⁷¹ in which the court enjoined the use of the name “Kids ‘r’ Us.” In *Toys “R” Us*, the court determined that for a survey to be trustworthy, the party presenting the survey must provide

evidence that (1) the ‘universe’ was properly defined, (2) a representative sample of that universe was selected, (3) the questions to be asked of interviewees were framed in a clear, precise and non-leading manner, (4) sound interview procedures were followed by competent interviewers who had no knowledge of the litigation or the purpose for which the survey was conducted, (5) the data gathered was accurately reported, (6) the data was analyzed in accordance with accepted statistical principles and (7) objectivity of the entire process was assured.⁷²

The court further explained that “[f]ailure to satisfy one or more of these criteria may lead to exclusion of the survey”⁷³ and that “[t]he necessary foundation will normally be laid through the

⁶⁹ For a discussion of control stimuli and related cases, see Edelman, *supra* note 64, at 405-17. See also Rapoport, *supra* note 62, at 987 (stating that if “a control stimulus is so different that no respondent would ever guess [an association with the plaintiff’s original product], even when it is the only stimulus shown, then such a control cannot measure whether people are guessing with response to the stimulus of actual interest . . .”).

⁷⁰ See *supra* note 22 and accompanying text for an explanation of why the Second Circuit specifically was chosen for this analysis. For a discussion of how the Ninth Circuit has treated some of the elements discussed in this article that does not focus on the question of survey stimuli, see Robert H. Thornburg, *Trademark Survey Evidence: Review of Current Trends in the Ninth Circuit*, 21 SANTA CLARA COMPUTER & HIGH TECH. L.J. 715 (2005).

⁷¹ *Toys “R” Us, Inc. v. Canarsie Kiddie Shop, Inc.*, 559 F. Supp. 1189 (E.D.N.Y. 1983).

⁷² *Id.* at 1205 (citations omitted).

⁷³ *Id.*

testimony of the persons responsible for the various parts of the survey.”⁷⁴ The subsequent trademark-related cases in the Second Circuit either directly cited the *Toys “R” Us* test or employed substantially similar factors that demanded a proper sample, survey content, and procedure. The issue of a researcher’s choice of the appropriate survey stimulus has received frequent scrutiny in the Second Circuit’s trademark jurisprudence and has sometimes had a significant impact on a court’s evaluation of a survey’s overall trustworthiness. To evaluate the court’s understanding of survey content and decide whether it has been consistent, it is most helpful to begin by dividing the main cases that have explicitly discussed survey stimuli into two groups: those that courts in the Second Circuit have deemed acceptable and those that they have not.

2. Survey Stimuli Accepted by the Second Circuit⁷⁵

a. Surveys Employing the Product Itself or an Image Thereof

Almost all of the surveys that courts in the Second Circuit have considered reliable in trademark infringement cases have used either the allegedly infringing product itself or an image of the infringing product as the stimulus that was presented to subjects. These cases involved a great diversity of products, and courts approved of surveys that, for example, showed pictures of various company signs including that of an allegedly infringing pharmaceutical company;⁷⁶ a photograph of an allegedly infringing revolver with all markings in plain view;⁷⁷ “a color photograph of . . . [an allegedly infringing] knife accompanied by a catalogue-styled heading that read ‘11-Function Swiss Army knife’ and was followed by a description of the knife”;⁷⁸ a line-up of

⁷⁴ *Id.*

⁷⁵ Some, but not all, of the surveys mentioned in this Subpart were used in jury trials. These surveys are all valuable in understanding judges’ reasoning concerning survey stimuli. This article, therefore, will not make the distinction for each case.

⁷⁶ *Sterling Drug, Inc. v. Bayer AG*, 14 F.3d 733 (2d Cir. 1994). In *Sterling Drug*, an American pharmaceutical company was suing a German pharmaceutical company for infringing on the “Bayer” trademark by creating several types of consumer confusion as to the affiliation between the two companies. The question accompanying the photographs in the survey asked “[w]hat type of product or products, if any, are made by the company or companies mentioned in the sign.” *Id.* at 741 (internal quotation marks omitted).

⁷⁷ *New Colt Holding Corp. v. RJG Holdings of Fla., Inc.*, 312 F. Supp. 2d 195 (D. Conn. 2004) (involving a trademark and trade dress infringement suit between a firearm manufacturer and its competitor). The court emphasized that the survey appropriately “focuses on direct perception of the marks, even if through photographs.” *Id.* at 224-25.

⁷⁸ *Forschner Group, Inc. v. Arrow Trading Co.*, 904 F. Supp. 1409, 1426-27 (S.D.N.Y. 1995). *Forschner* pitted an importer of pocket knives from Switzerland, designated as “Swiss Army Knives,” against a distributor of knives made in China, who was accused of misrepresentation of origin and false advertising.

six cans of condensed milk including the plaintiff's product and the defendant's competing and allegedly infringing product;⁷⁹ a popcorn product sold in champagne bottles alleged to infringe on a champagne manufacturer's trademark;⁸⁰ a plaintiff's nail polish bottle in a survey seeking to demonstrate that the bottle design had acquired secondary meaning;⁸¹ garments carrying defendant's allegedly infringing logo;⁸² a photograph of defendant's blue "Metbank" logo and a similar blue logo with the control mark "MetroBank";⁸³ a picture of the defendant's cigar products;⁸⁴ images of jeans on videotape, including allegedly infringing jeans as eliciting probative value to show the similarity of the litigating parties' jeans-stitching patterns.⁸⁵ Finally, the court in *Lon Tai Shing Co. v. Koch + Lowy*⁸⁶ permitted a survey in which a "mystery shopper" went to lamp stores and showed salesmen one of three different photographs—the plaintiff's lamp, the defendant's lamp, or a control lamp.⁸⁷ The court explained that the advantage of this type of study "is that it most closely approximates real market conditions."⁸⁸

⁷⁹ *Friesland Brands, B.V. v. Vietnam Nat'l Milk Co.*, 221 F. Supp. 2d 457, 459-61 (S.D.N.Y. 2002).

⁸⁰ *Schieffelin & Co. v. Jack Co. of Boca, Inc.*, 850 F. Supp. 232, 247 (S.D.N.Y. 1994) (allowing survey evidence of actual confusion between the plaintiff's trademark "Cuvee Dom Perignon" and the popcorn "Dom Poppingnon" sold in champagne bottles with a similar label).

⁸¹ *Essie Cosmetics, Ltd. v. Dae Do Int'l, Ltd.*, 808 F. Supp. 952, 955 (E.D.N.Y. 1992).

⁸² *PRL U.S.A. Holdings, Inc. v. U.S. Polo Ass'n, Inc.*, No. 99-CV-10199, 2006 WL 1881744, at *1 (S.D.N.Y. July 7, 2006).

⁸³ *MetLife, Inc. v. Metro. Nat'l Bank*, 388 F. Supp. 2d 223, 232-33 (S.D.N.Y. 2005). The survey found that significantly more people believed that plaintiff Metlife owned or sponsored "MetBank" (38%) than the control company "MetroBank" (4%). *Id.* at 233.

⁸⁴ *Empresa Cubana Del Tabaco v. Culbro Corp.*, No. 97 Civ. 8399, 2004 WL 602295, at *26 (S.D.N.Y. Mar. 26, 2004), *rev'd in part*, 399 F.3d 462 (2d Cir. 2005), *cert. denied*, 126 S. Ct. 2887 (2006).

⁸⁵ *Lois Sportswear, U.S.A., Inc. v. Levi Strauss & Co.*, 799 F.2d 867, 869 (E.D.N.Y. 1986).

⁸⁶ *Lon Tai Shing Co. v. Koch + Lowy*, No. 90 Civ. 4464, 1991 WL 170734 (S.D.N.Y. June 20, 1991).

⁸⁷ *Id.* at *17-18.

⁸⁸ *Id.* at *19. A number of surveys have been rejected on grounds other than their stimuli. Thus, while it is uncertain whether they would have been admitted had it not been for other flaws, these surveys provide further evidence of the fact that trademark surveys tend to use visual stimuli. *See, e.g.*, *Starter Corp. v. Converse*, 170 F.3d 286, 296-97 (2d Cir. 1999) (using five brands of athletic shoes as stimuli); *Mastercard Int'l Inc. v. First Nat'l Bank of Omaha, Inc.*, No. 02 Civ. 3691, 2004 WL 326708, at *7-9 (S.D.N.Y. Feb. 23, 2004) (where a stimulus consisting of "a general information page, a description of the smart card program, and a prototype of a 'OneSmart Mastercard'" in a survey, sought to show that the ONESMART mark would not be confused with the SMART ONE mark); *Medical Econ. Co. v. Prescribing Reference, Inc.*, 294 F. Supp. 2d 456, 461 (S.D.N.Y. 2003) (stimuli consisting of the plaintiff's, the defendant's, and a control actual publication product); *Jordache Enters., Inc. v. Levi Strauss & Co.*, 841 F. Supp. 506, 518-19 (S.D.N.Y. 1993) (allegedly infringing jeans with all labelings as stimulus); *Weight Watchers Int'l, Inc. v. Stouffer Corp.*, 744 F. Supp. 1259, 1272-76 (S.D.N.Y. 1990) (allegedly infringing print advertisement used as stimulus in both plaintiff's and defendant's surveys); *Dreyfus Fund, Inc. v. Royal Bank of Canada*, 525 F. Supp. 1108, 1116 n.2 (S.D.N.Y. 1981) (stimuli

In all of these cases, researchers showed subjects either the product itself or a picture that displayed the product in an unaltered condition. The way in which the stimulus was presented, as well as the setting in which it was shown, varied from study to study. Courts thus gave a lot of latitude to researchers in how to define the market conditions external to the stimulus itself: the *Lon Tai Shing* court approved of researchers' decision to conduct surveys in stores themselves, but this was by no means a requirement for the manner in which to show a stimulus. It appears that courts have placed great value on the stimulus choice itself by using it as such a strong proxy of market conditions. This judicial attitude appears rather sensible at first blush: after all, people will encounter a particular product in a variety of settings, and it is thus more important whether a stimulus truly represents a product than what the particular surroundings of a study happen to be. It is a bit worrisome, however, to accept as a given that the surroundings must have been "representative enough." As will become apparent in Part III.A.2, courts have set some limits to the environmental cues that subjects may receive while viewing stimuli. It remains unclear in many cases that have included judicially approved surveys, however, whether courts examined the issue of external cues at all.

b. The Great Exception: The *McBagel's* Phone Survey

The main case that admitted a survey despite its lack of a stimulus consisting of or resembling the product is *McDonald's Corp. v. McBagel's, Inc.*⁸⁹ In *McBagel's*, the court accepted a phone survey through which the plaintiff, McDonald's, sought to establish trademark infringement.⁹⁰ The defendants criticized the survey for "ignoring market conditions, since it failed to include a reference to defendant's 'bagel bakery restaurant' as advertised in print with the name 'McBAGEL's', or to the different typeface defendants use in print advertising."⁹¹ The court responded that "[a]lthough the survey did not replicate the parties' print advertising, by soliciting audio responses it was closely related to the radio advertising involved in the case."⁹²

The *McBagel's* case has little in common with the previously discussed ones. Its study's failure to present subjects with visual stimuli creates several problems. If one considers the average

consisting of corporate symbols, including the allegedly infringing lion symbol).

⁸⁹ *McDonald's Corp. v. McBagel's, Inc.*, 649 F. Supp. 1268 (S.D.N.Y. 1986).

⁹⁰ *Id.* at 1277-78.

⁹¹ *Id.*

⁹² *Id.* at 1278.

person's strong reliance on vision in his daily life, one cannot help but wonder whether a customer standing in front of a McBagel's bakery and viewing its sign, which purportedly did not resemble McDonald's sign, would still be confused enough to enter the bakery, mistaking it for a McDonald's affiliate.⁹³ Refusing to display a visual stimulus also exacerbates the problem of external cues because researchers do not know what their study subjects happened to be viewing at the time—it could have been a wall, a McDonald's advertisement, or any number of other things. The judge's explanation concerning the link between radio advertising and the telephone study is unconvincing and provides a lot of confusion for prospective researchers by blurring the boundaries of the definition of the marketplace. The situation might have presented itself differently if subjects had heard McBagel's radio advertisements and then ordered McBagel's products over the phone. In that case, only auditory stimuli and communication would have existed throughout, with no other factors offsetting the potential confusion from auditory stimuli.⁹⁴

3. Survey Stimuli Rejected by the Second Circuit

a. Improper Stimuli and Improper Questions

A large number of studies have been rejected for failing to employ an appropriate stimulus, or, as a related matter, proper questions. In the often-cited *Universal City Studios, Inc. v. Nintendo Co.* case,⁹⁵ a telephone survey failed, *inter alia*, to “connect Donkey

⁹³ It is not necessarily the case that there would even have been so-called “initial interest” confusion for a hypothetical customer standing in front of a McBagel's bakery.

⁹⁴ Over the years, several courts have had to decide trademark infringement cases relating to the “McDonald's” trademark. McDonald's has generally succeeded in winning such litigations regarding products related to foods, as well as those related to other products. *See, e.g.*, *J & J Snack Foods Corp. v. McDonald's Corp.*, 932 F.2d 1460 (Fed. Cir. 1991) (disallowing registration of the trademarks “McPretzel” and “McDugal McPretzel” due to likelihood of confusion with the family of marks owned by McDonald's); *McDonald's Corp. v. Druck & Gerner, DDS., P.C.*, 814 F. Supp. 1127 (N.D.N.Y. 1993) (enjoining defendant dentists from using the name “McDental”); *Quality Inns Int'l v. McDonald's Corp.*, 695 F. Supp. 198, 211 (D. Md. 1988) (stating that “McDonald's golden arches and the McDonald's logo rank among the strongest marks, enjoying instant recognition among virtually all members of our society” and enjoining the use of the name “McSleep”); *McDonald's Corp. v. McKinley*, 1989 WL 274414, at *6 (TTAB Sept. 25, 1989) (holding that “applicant's use of MCTEDDY for teddy bears is likely to cause confusion as to source in view of opposer's long use of its heavily advertised and consequently well-known ‘MC’ and ‘MAC’ ‘family of marks’ on a variety of goods, including toys.”). Courts' suspicion of product names beginning with “Mc” in these cases may suggest that the *McBagel's* court gave a lot of weight to the survey in part because of the generally strong case that McDonald's is perceived to have in the trademark context.

In the author's view, one cannot reconcile the outcome of *McBagel's* with the rest of Second Circuit doctrine and with common principles of survey construction. The court should thus strike it down as bad law when a similar issue is next reviewed.

⁹⁵ *Universal City Studios, Inc. v. Nintendo Co.*, 746 F.2d 112 (2d Cir. 1984).

Kong with Universal's alleged trademark in the *name* King Kong" when it just asked participants: "To the best of your knowledge, was the Donkey Kong game made with the approval or under the authority of the people who produce the King Kong movies?"⁹⁶ The question was directly leading by suggesting a link between "Donkey Kong" and "King Kong" rather than allowing subjects to elicit answers themselves through open-ended questions such as "Who do you think created the game Donkey Kong?" The lack of a visual stimulus compounds the use of the leading question by pushing subjects to respond in a desirably biased way, and thus even high response rates, which generally indicate potential confusion, are unconvincing. The *Universal City Studio* case, however, did not reject phone surveys outright.

b. Lack of a Visual Stimulus

Courts have rejected surveys for failing to use stimuli: "such as pictures, advertisements or clothing, that directly expose potential consumers to the products or the marks in question";⁹⁷ "that would inform consumers as to the competing products or marks in question";⁹⁸ that should consist of "pictures or advertisements that

⁹⁶ *Id.* at 118. Unlike the court in *Universal City Studios*, the court in *Smithkline Beckman Corp. v. Proctor & Gamble Co.*, 591 F. Supp. 1229 (N.D.N.Y. 1984), authorized the use of a phone survey, but its sole purpose was to study consumer awareness of the brand name "Ecotrin," and *not* to establish infringement. *Id.* at 1237-38.

⁹⁷ *Troublé v. Wet Seal, Inc.*, 179 F. Supp. 2d 291, 308 (S.D.N.Y. 2001) (concerning an infringement lawsuit by clothes seller and owner of the trademark "Agnès b." against the owner of the trademark "Arden B") (citations omitted). The court also criticized the questions used in the survey:

The questions in the surveys included questions such as, "Do you think that 'Arden B.' clothing is made and sold by a company that makes or sells any other line of clothing?" and "Do you think the company that makes and sells 'Arden B.' clothing is affiliated, associated or connected with any other women's clothing company?" Rather than test a consumer's confusion, these question[s] test their awareness of other lines of clothing and company affiliation. If an interviewee spontaneously answered 'Agnès b.' to the second question, there would be some evidence of confusion. But contrary to the assertions the survey's author . . . , the fact that the majority of interviewees answered "no" to the second question does little to establish that they would not be confused if they encountered both marks.

Id. at 308 n.12. On a related note, the same court found a survey unacceptable that presented subjects with a blouse and skirt with the "Notorious" trademark as well as a fragrance bottle and its carton with the allegedly infringing mark of the same name, and asked the participants if they thought the two products were made by the same or a different company; the court specifically criticized that "the side-by-side alignment of the products is a somewhat unrealistic simulation of actual market conditions." *Edison Bros. Stores Inc. v. Cosmair Inc.*, 651 F. Supp. 1547, 1558-59 (S.D.N.Y. 1987).

⁹⁸ *1-800 Contacts, Inc. v. WhenU.com*, 309 F. Supp. 2d 467, 500 (S.D.N.Y. 2003) (involving a suit in which the owner of a website and mark sought to enjoin a competitor and a software company from delivering competitive pop-up advertisements to computer users), *rev'd on other grounds*, 414 F.3d 400 (2d Cir. 2005), *cert. denied*, 126 S. Ct. 749. The decision mentioned that "[t]ypically, trademark infringement surveys use stimuli, such as pictures, advertisements or clothing, that directly expose potential consumers to the products or the marks in question." *Id.* (quoting *Wet Seal*, 179 F. Supp. 2d at 308).

approximate what a potential customer would encounter in making her television-viewing choices”;⁹⁹ that would correspond to actual products containing the disputed term rather than just a card with the term written on it,¹⁰⁰ and that would present a visual rather than just aural representation of the covers of magazines alleged to display infringing material.¹⁰¹ In the recent case

Additionally, the court criticized some of the questions used in the survey:

Question 4-2 reads, “I believe that anyone should have the right to place ‘Pop-Up’ advertisements on any website at any time, even if the owner of the website does not authorize or approve it” By suggesting in the second clause that the pop-up ads might be unauthorized, [the] survey suggests that they should not be permitted on the website. Question 4-5 reads: “I believe that ‘Pop-Up’ advertisements are sometimes not sponsored by or authorized by the website on which they appear.” However, Question Nine reads: “Were you aware that, when viewing websites on the Internet, SaveNow software causes ‘Pop-Up’ advertisements to be displayed on your computer which are not authorized by the website on which they appear?” Since Question Nine flatly states that pop-up ads generated by SaveNow software are unauthorized, the survey itself suggests the answer to Question 4-5.

1-800 Contacts, Inc., 309 F. Supp. 2d at 500 n.54 (citations omitted). The court concluded that the survey did not provide evidence of actual confusion and was of “weak probative value in establishing the likelihood of initial interest confusion,” which “weighs in favor of neither Defendant nor Plaintiff.” *Id.* at 501.

⁹⁹ *WE Media, Inc. v. Gen. Elec. Co.*, 218 F. Supp. 2d 463, 474 (S.D.N.Y. 2002) (involving a suit where a media company sued cable television channel producers for trademark infringement and dilution).

¹⁰⁰ *Juicy Couture, Inc. v. L’Oreal U.S.A., Inc.*, No. 04 Civ. 7203, 2006 WL 1012939, at *25 (S.D.N.Y. Apr. 19, 2006) (involving a suit by the owner of marks such as “Juicy,” “Juicy Couture,” “Choose Juicy,” etc. and the allegedly infringing owner of “Juicy Wear,” “Juicy Pop,” and “Juicy Pop Gossip”).

¹⁰¹ *Inc. Publ’g Corp. v. Manhattan Magazine, Inc.*, 616 F. Supp. 370, 392-93 (S.D.N.Y. 1985) (pitting the publisher of a business magazine against a competitor in a trademark infringement suit). Specifically, the court stated:

The respondents in a telephone survey are responding only to aural stimuli; in layman’s language, to the sounds of names. But we are concerned with magazines. The covers of magazines are seen. The contents of magazines are read. Such activities give rise to visual and cerebral stimuli. They play no part in a telephone survey such as the one at bar.

Id. at 393; *accord* *Penta Hotels Ltd. v. Penta Tours*, 9 U.S.P.Q.2d 1081, 1104-05 (D. Conn. 1988) (dismissing a survey for its suggestive question: “Do you feel that there is some connection between Penta Hotels and Penta Tours or do you feel that there is no connection between these companies[.]” and stating that “[t]he use of a telephone in conducting the survey here does not make it inherently deficient, although the court recognizes that respondents were unable to view the parties’ logos through the telephone and thus the survey is not valid on the issue of similarity of the logos.”). The court in *Manhattan Magazine* also criticized the questions used in the survey:

Under question 3a, the interviewer says this over the telephone to the respondent: Now I’m going to read the names of five magazines. Please listen to the names of these magazines and tell me if any of these magazines sound to you as if they might have the same publisher.” The interviewer then read the names of the following five magazines: “Business Week”; “Fortune”; “Inc.”; “Money Magazine”; and “Manhattan, inc.” The interviewer then asked: “Now, do any of the magazines I just mentioned sound as if they might have the same publisher?” The respondent was asked to respond “yes” or “no.” The interviewer then moved to question 3b, and said: “Which magazines sound to you as if they might have the same publisher?” The interviewer then posed question 3c: “Why do the magazines you mentioned sound as if they might have the same publisher? On question 4a, the interviewer said to the respondent: “I’m going to read the list of magazines one more time. This time please tell me if any of these magazines sound to you as if they might contain similar types of

*Starbucks Corp. v. Wolfe's Borough Coffee, Inc.*¹⁰² mentioned in the Introduction,¹⁰³ the plaintiff conducted a phone survey in which it asked people to “name any company or store that [they] think might offer a product called ‘Charbucks.’”¹⁰⁴ The court explained that, in addition to the fact that only a small percentage of individuals named ‘Starbucks’ as the answer, “in light of the fact that the survey item was administered by telephone and did not present the term ‘Charbucks’ in the linguistic or visual context which Defendant uses it, this evidence is insufficient to make the actual confusion factor weigh in Plaintiff’s favor to any significant degree.”¹⁰⁵

Standing alone, the above-mentioned examples would seem to suggest fairly clearly that visual stimuli are required, but the Second Circuit’s lack of complete bans on phone surveys prevents one from drawing this conclusion. Note that phone surveys are not always conducted with nefarious motives of eliciting bias: they are often simply less expensive and faster than in-person interviews. With the advent of the Internet, however, their justifiability is likely to decrease. Internet studies are able to present both accurate pictures and sounds and prevent the survey noise that can always arise from the particular interaction between a researcher and a subject (whether that interaction happens on the phone or in person). Internet surveys will, however, presumably still present problems in some cases due to their remoteness from the environment in which a product sale would occur.¹⁰⁶

c. Visual Stimuli Altered from the Original

In *American Footwear Corp. v. General Footwear Co.*,¹⁰⁷ an older case that involved a dispute about the trademark “Bionic,” the

articles and information” For question 5, the interviewer asked the respondent the following: “Do you think Manhattan, inc. magazine might be a regional version of Inc. magazine?” The three recorded answers were “yes,” “no,” or “don’t know.” The questionnaire then concluded with eleven questions for “classification purposes,” dealing with work, work title, education, age, income and the like.

Manhattan Magazine, 616 F. Supp. at 391.

¹⁰² *Starbucks Corp. v. Wolfe’s Borough Coffee, Inc.*, No. 01-Civ-5981, 2005 WL 3527126 (S.D.N.Y. Dec. 23, 2005).

¹⁰³ See *supra* Part I.

¹⁰⁴ *Starbucks Corp.*, 2005 WL 3527126, at *5 (citation omitted).

¹⁰⁵ *Id.* It is not entirely clear from the judge’s formulation whether he decided to admit the survey and did not lend it much weight or whether he did not even think it should be admitted. This may, however, largely constitute a semantic difference without much consequence as to the outcome.

¹⁰⁶ For a discussion of when Internet surveys would be appropriate, see *infra* Part IV.

¹⁰⁷ *Am. Footwear Corp. v. Gen. Footwear Co.*, 609 F.2d 655 (2d Cir. 1979).

court rejected two surveys for improper questions and stimuli.¹⁰⁸ The first survey was a telephone study that asked individuals: “With whom or what do you associate a product labelled Bionic?” According to the court, the survey should have asked: “With whom or what do you associate a ‘Bionic’ boot?”¹⁰⁹ The second survey showed participants an American promotional poster for the “Bionic Boot” and asked them: “With whom or what do you associate the term Bionic Boot?” The court rejected this survey because the poster had been modified from the original poster displayed at various shoe fairs in that references to American Footwear as the seller of the boot had been removed.¹¹⁰ The court raised similar concerns in *Jaret International, Inc. v. Promotion in Motion*,¹¹¹ where a manufacturer of sour candy accused a competitor of trade dress infringement.¹¹² The court rejected the plaintiff’s survey, because, among other reasons, “there is no dispute that either SOUR PATCH KIDS or SOUR JACKS purchasers would ever purchase these products without the full marks on the boxes, the survey technique of removing portions of them is suspect.”¹¹³ The court came to the same conclusion in a case involving two food distributors in which the plaintiff’s survey did not use plaintiff’s complete trademark “Vista Food Exchange, Inc.” with its corresponding flag logo but rather only the word “Vista” in the brochure that survey respondents viewed before they were asked what company produced the brochure.¹¹⁴ Finally, in *Revlon Consumer Products Corp. v. Jennifer Leather Broadway, Inc.*,¹¹⁵ the court criticized the fact that the plaintiff’s survey “only tested consumers’ reactions to the black and white newspaper version of defendant’s advertisement. [The plaintiff’s survey] made no attempt to study the subway version of the advertisement which depicted the couches in nine different colors.”¹¹⁶

It is interesting to see that the Second Circuit deems the use of imperfect visual stimuli in these cases as worse than using no visual stimulus at all, such as in a phone study. Removing significant portions of visual stimuli is thus never acceptable,

¹⁰⁸ *Id.* at 660.

¹⁰⁹ *Id.*

¹¹⁰ *Id.*

¹¹¹ *Jaret Int’l, Inc. v. Promotion in Motion, Inc.*, 826 F. Supp. 69 (E.D.N.Y. 1993).

¹¹² *Id.* at 73-74.

¹¹³ *Id.* (stating that the survey “has the same methodological defect as the one cited in *American Footwear*”).

¹¹⁴ *Vista Food Exch., Inc. v. Vistar Corp.*, No. 03-CV-5203, 2005 WL 2371958, at *5 (E.D.N.Y. Sept. 27, 2005).

¹¹⁵ *Revlon Consumer Prods. Corp. v. Jennifer Leather Broadway, Inc.*, 858 F. Supp. 1268 (S.D.N.Y. 1994).

¹¹⁶ *Id.* at 1275. The court also mentions that one of the questions in the survey was: “Do you think Revlon is in any way associated with Jennifer Leather?” *Id.* at 1276.

because it becomes a selective tool that manipulates and biases subjects. The underlying assumption is that some aspects of a product may resemble another product, but that the resemblances are offset by remarkable differences that cannot simply be eliminated in a survey to produce a desired result. A truly infringing product is one that elicits confusion in the actual marketplace, which means that the similarities with another product are so great that none of the variations make up for it.

d. Inappropriate Visual Control Stimuli

In *Conopco, Inc. v. Cosmair, Inc.*,¹¹⁷ the owner of a trademarked perfume bottle alleged infringement and dilution by its competitor. The court rejected a survey that the plaintiff conducted using the defendant's *pre-launch* bottle as a stimulus rather than the one that consumers would eventually see in the marketplace, and explained that "a survey must use the proper stimulus, one that tests for confusion by replicating marketplace conditions."¹¹⁸ This language was borrowed from *Cumberland Packing Corp. v. Monsanto Co.*,¹¹⁹ which held that a manufacturer of artificial sweetener did not infringe upon its competitor's trademark.¹²⁰ In a later decision, the court in *Cumberland Packing Corp. v. Monsanto Co.*,¹²¹ rejected a revised survey presented by the plaintiff in part because it used improper, artificial-looking control stimuli.¹²² Similar to original stimuli that have been altered,

¹¹⁷ *Conopco, Inc. v. Cosmair, Inc.*, 49 F. Supp. 2d 242 (S.D.N.Y. 1999).

¹¹⁸ *Id.* at 253-54.

¹¹⁹ *Cumberland Packing Corp. v. Monsanto Co.*, 32 F. Supp. 2d 561 (E.D.N.Y. 1999) [hereinafter *Cumberland*, 1999].

¹²⁰ *Id.* at 575.

¹²¹ *Cumberland Packing Corp. v. Monsanto Co.*, 140 F. Supp. 2d 241 (E.D.N.Y. 2001) [hereinafter *Cumberland*, 2001].

¹²² Specifically, the court stated:

The variations of the NutraSweet box, both light blue and the darker blue, with the coffee cup removed are inappropriate and thoroughly unrealistic. When the coffee cup was removed it was not replaced with any other graphic and every other detail on the box was unchanged. In place of the coffee cup was a conspicuous absence in the middle of the front panel. The sweetener packet, tilted on angle so as to lean against a cup, was left tilting in a seemingly random fashion. The results of the survey indicate that these two non-coffee cup variations caused the lowest levels of confusion. This is hardly surprising. These boxes look unauthentic and contrived. Comparing the confusion levels caused by these two so-called controls to the other NutraSweet box variations cannot result in any legally meaningful data.

Id. at 249. The *Cumberland*, 2001 court also stated: "The test is not whether the products can or do create confusion in the abstract but whether that confusion affects the purchasing and selling of the goods or services in question." *Id.* at 250 (citations omitted). Survey subjects were asked the following questions:

Q1. Please look at this box of aspartame-based sugar substitute as if you were considering buying it. Q2. Do you think that any of these sugar substitutes are made by the same company that made the first box you saw? Q3. Which one or more of these sugar substitutes do you think are made by the same company as

inappropriate control stimuli create bias by suggesting a particular answer that may not correlate with confusion in the marketplace. The *Cumberland* court also implicitly seems to have asked itself whether a researcher with no specific result in mind would have ever produced that survey. The unnatural displays of the control sweetener packets in the survey made the court answer this question in the negative. The same question was asked and answered in the negative by the courts that decided the cases in which visual stimuli were altered from the original by removing product packaging markings and the like.¹²³ In another case, the court found that the plaintiff owner of “Ice Breakers,” who was suing the owner of “Dentyne Ice,” should have employed a control product other than “Trident” because “[a]lthough Trident is a gum product, it fails to capture the essence of the allegedly confusing quality at issue, namely the ‘Ice’ term or some variation of that theme.”¹²⁴

e. Failure to Present Both the Original and Allegedly Infringing Product

Courts have generally been unsympathetic toward parties that presented surveys that did not utilize both the plaintiff’s and the defendant’s product. In one case, the defendant’s expert only presented survey respondents with the defendant’s literary products rather than displaying the plaintiff’s goods as well, which the court found could have greatly underestimated the likelihood of consumer confusion.¹²⁵ In another instance, the court held that a defendant’s survey was flawed because it failed to display the plaintiff’s alcoholic product in addition to its own.¹²⁶ Thus, it is safe to say that a survey seeking admission in the Second Circuit should present both the plaintiff’s and the defendant’s marks.

the first box you saw? Q4. Why do you say that? Q5. If the respondent mentioned color as a reason for thinking that one of the sugar substitutes in the second room was made by the same company that made the first box they saw they were asked: When you say color or coloring, precisely what do you mean by that?

Id. at 246-47.

¹²³ See *supra* Part III.A.3.d.

¹²⁴ *Nabisco v. Warner-Lambert Co.*, 32 F. Supp. 2d 690, 700 (S.D.N.Y. 1999). The court also criticized the survey at bar for using advertisements for “Ice Breaker” and “Dentyne Ice” rather than the products themselves. *Id.*

¹²⁵ *Simon & Schuster, Inc. v. Dove Audio, Inc.*, 970 F. Supp. 279, 291 (S.D.N.Y. 1997) (pitting the owner of the trademark “The Book of Virtues” against the maker of “The Children’s Books of Virtues” products).

¹²⁶ *Star Indus., Inc. v. Bacardi U.S.A., Inc.*, No. 02 Civ. 4239, 2003 U.S. Dist. LEXIS 23411, at *12 (S.D.N.Y. Dec. 31, 2003) (citing *Simon & Schuster, Inc.*, 970 F. Supp. at 281 and holding in favor of the defendants despite their flawed survey evidence because the plaintiff failed to produce any survey), *aff’d*, 412 F.3d 373 (2d Cir. 2005).

f. Compound Problems

A number of problems related to survey methodology surfaced in *Pilot Corp. v. Fisher-Price, Inc.*,¹²⁷ where the plaintiff sued its former licensee for infringing upon the mark of the drawing toy “Magna Doodle” by using the term “Doodle Pro.” The court criticized the plaintiff’s survey for failing to show the actual products rather than just their logos accompanied by a brief description of the toys they designate; the survey only showed parts of the Magna Doodle logo (without the words “The Original,” and the Magna Doodle Slogan, that accompany both Magna Doodle products); and the survey failed to employ a proper control stimulus by only using purple writing on a yellow background to designate the two products in dispute without any of the control items.¹²⁸ For these reasons, the court gave more weight to the defendant’s two surveys that did not contain these flaws.¹²⁹

The plaintiff and defendant in *National Distillers Products Co. v. Refreshment Brands, Inc.*,¹³⁰ each presented a survey to determine whether there was infringement of the “Teton Glacier” mark associated with high quality vodka. The court rejected the plaintiff’s survey, mentioning the following among its methodological errors:

First, the respondents were shown the Teton Glacier label on a card, rather [than] on a bottle, as they would encounter the label in the marketplace. Second, while the cooler bottles would obviously be full when sold in the marketplace, they were empty when shown to the subjects of the [plaintiff’s] Survey The [defendant’s] Survey, on the other hand, is much more reliable because it more accurately approximates actual market conditions. For example, subjects were not made artificially aware of . . . the TETON GLACIER trademark.¹³¹

The defendant’s survey, which employed a bottle and a four-pack of the allegedly infringing product, and a bottle and four-pack of a

¹²⁷ *Pilot Corp. v. Fisher-Price Inc.*, 344 F. Supp. 2d 349 (D. Conn. 2004).

¹²⁸ *Id.* at 358.

¹²⁹ *Id.* at 358-59.

¹³⁰ *Nat’l Distillers Prods. Co. v. Refreshment Brands*, 198 F. Supp. 2d 474 (S.D.N.Y. 2002). The questioning of subjects in the plaintiff’s survey proceeded as follows:

In the first “room,” the interviewer showed the subject the TETON GLACIER label, without the bottle to which it is normally attached. Then, in the second room, the subject was shown seven or eight empty cooler bottles and told, “Some, all, or none of these alcoholic beverages may come from the same maker or company as the product on the card I showed you [in the first room]. Which, if any, of these alcoholic beverages do you think come from the same maker or company as the product on the card I showed you [in the first room]?”

Id. at 482-83 (footnote omitted).

¹³¹ *Id.* at 484 (footnote omitted).

control product, was thus found to be acceptable.¹³² While the alteration of original visual stimuli has been discussed,¹³³ the Teton Glacier label presents a slightly different issue. Nobody claimed that the label was altered, but only that it was presented on a card rather than a bottle. In a number of cases that either accepted visual stimuli¹³⁴ or mentioned that they should have been present,¹³⁵ courts did not seem to differentiate between a product itself and an accurate image thereof. One therefore puzzles over whether the Teton Glacier card was problematic because it only represented an *image* rather than a *product*, or rather because it displayed the image of the *label* rather than of the entire *bottle*. The former would not fit well with the other case law in the Second Circuit; the latter is problematic as well because it is unclear whether the problem is an issue of *ex ante* principle or rather one of an *ex post* judgment made after having seen the results of the defendant's study. This case illustrates, once more, the problem of not providing clearer guidelines as to what types of stimuli are and are not acceptable.

g. Environmental Cues

Courts have also paid attention to cues in the environment in which surveys are conducted. In one case,¹³⁶ the district court criticized a plaintiff's survey because "large signs bearing plaintiff's logo were visible from some of the test sites or on the way to some of the test sites."¹³⁷ In *Hutchinson v. Essence Communications, Inc.*,¹³⁸ a rap music performer sued for declaration of non-infringement of trademark for use of the name "Essence" against the publisher of *Essence* magazine. The court refused to give significant weight to a survey that the defendant commissioned.¹³⁹ One of the criticisms mentioned in the case is that survey participants could

¹³² *Id.* at 483-84. The defendant's survey used the following questioning:

The subjects were shown either the Arctic Bay or Glacier Bay products and permitted to examine them for as long as they wished. The products were then taken away and the subjects were asked a series of questions, to wit: "What company or companies do you think makes or puts out this product?"; "Whether you know the name of the company that makes this product, are you aware of any other products or brands put out by this company?"; if yes, then, "What other products or brands do you think are put out by the company that makes this product?"; and "If you have an opinion, do you think that the company that makes this product did or did not get approval or permission from any other company or companies in order to put out this product?"

Id. at 483.

¹³³ *See supra* Part III.A.3.d.

¹³⁴ *See supra* Part III.A.2.a.

¹³⁵ *See supra* Part III.A.3.c.

¹³⁶ *Exxon Corp. v. XOIL Energy Res., Inc.*, 552 F. Supp. 1008 (S.D.N.Y. 1981).

¹³⁷ *Id.* at 1022.

¹³⁸ *Hutchinson v. Essence Commc'ns, Inc.*, 769 F. Supp. 541 (S.D.N.Y. 1991).

¹³⁹ *Id.* at 557-65.

possibly see a copy of *Essence* magazine during the interview, which may have influenced their answers.¹⁴⁰ Of further interest is the *Hutchinson* court's explanation of why *Hutchinson* should be distinguished from *McBagel's*. Notably, the court explains that "restaurants are often franchises; singers rarely are. Thus, it is more likely that the question did not plant an idea in the minds of the interviewees in the restaurant context than in this context, given the judicially noticeable prevalence of restaurant chains."¹⁴¹ This explanation is far from convincing, especially because it does not seem to stem from any empirical evidence. It showed, however, that the court was unwilling to overrule *McBagel's* and thus still considered it good law.

h. Leading Brands

In *Mennen Co. v. Gillette Co.*,¹⁴² a manufacturer of deodorants and antiperspirants sought to prevent the defendant competitor from using a diagonal stripe design on its packaging. The court rejected the use of a survey said to be "fatally defective and untrustworthy in a number of ways" and explained:

In this survey, respondents were shown the Mennen stripes and asked: "Please look at this picture and this list of stick deodorant brand names. Among the leading brands of stick deodorants now on the market, which of the brands listed here would you say uses these stripes on their package?" The respondents were then shown a list of manufacturers. While 63% of the respondents selected Mennen from the array, this was merely a playback of brand share in response to a highly suggestive, aided question, in which respondents most likely pointed to Mennen as it was the major producer of stick deodorants.¹⁴³

Similarly, in *Pfizer Inc. v. Astra Pharmaceutical Prods., Inc.*,¹⁴⁴ the court rejected a finding of secondary meaning on the basis of a survey that sought to establish that defendant's drug TOPROL XL infringed on plaintiff's mark PROCARDIA XL. In that case, a survey asked: "Thinking of the practice of cardiovascular medicine, what first comes to mind when you hear the letters XL?"¹⁴⁵ The court explained:

It is not probative of secondary meaning that physicians conditioned (or asked) to respond with a drug name

¹⁴⁰ *Id.* at 565.

¹⁴¹ *Id.*

¹⁴² *Mennen Co. v. Gillette Co.*, 565 F. Supp. 648 (S.D.N.Y. 1983).

¹⁴³ *Id.* at 652.

¹⁴⁴ *Pfizer Inc. v. Astra Pharm. Prods., Inc.*, 858 F. Supp. 1305 (S.D.N.Y. 1994).

¹⁴⁵ *Id.* at 1321.

responded with one of the most popular, if not the most popular, cardiovascular drug on the market and the only XL-suffixed brand having significant presence in the market at the time.¹⁴⁶

Why this logic was not also used in *McBage*'s is unclear. Generally, such reasoning appears to prevent large brands from artificially entrenching themselves by using their status difference in leading questions. It has been suggested that certain types of survey questions, including the type that require subjects to identify a company that makes a potentially infringing product, overestimate “the likelihood of consumer confusion about the sources of products and the extent of consumer confusion that can validly be attributed to similarity in the names, sounds, and appearances of consumer goods or services” and show “that the demand effects of obtrusive questioning create major response bias and distortion that favor the well-known brand or company.”¹⁴⁷

This category of surveys particularly highlights the fundamental problem of failing to provide *ex ante* standards for stimuli and questions. A lot of money must have been spent on conducting these surveys—whether the researchers intentionally usurped the big brand status of their clients is not entirely clear—and the whole issue could have been avoided had the court provided better guidelines beforehand.¹⁴⁸ Even after these cases, future researchers may only have a few more leads as to what types of questions to avoid when dealing with major brands, but nothing will prevent them from making slightly different mistakes.

i. Business Associations Through Similar Names

In *Beneficial Corp. v. Beneficial Capital Corp.*,¹⁴⁹ the plaintiff claimed that “defendants’ use of the name ‘Beneficial’ is likely to cause confusion as to the source of defendants’ services and to induce the public to deal with defendants in the mistaken belief that the services offered by defendants are in fact those of plaintiffs.”¹⁵⁰ The court stated:

To the key question asked by the survey, “Do you think that there may or may not be a business connection between

¹⁴⁶ *Id.* at 1321-22.

¹⁴⁷ John P. Liefeld, *How Surveys Overestimate the Likelihood of Consumer Confusion*, 93 TRADEMARK REP. 939, 941 (2003). *Cf.* A. David Morrow & Ruth M. Corbin, *Pulling Confusion Surveys Back from an Illusory Brink: Reply to an Article of Dr. John Liefeld*, 94 TRADEMARK REP. 1372, 1380-81 (2004) (disagreeing with Liefeld and claiming that the bias toward better-known brands can be corrected through proper control conditions).

¹⁴⁸ Such problems could have similarly been avoided if a court had declared the survey improper on its face, as will be discussed *infra* Part IV.

¹⁴⁹ *Beneficial Corp. v. Beneficial Capital Corp.*, 529 F. Supp. 445 (S.D.N.Y. 1982).

¹⁵⁰ *Id.* at 447.

Beneficial Capital Corp. and the Beneficial Finance System Companies?” thirty-one percent of the respondents stated that such a connection was either definite or probable. This evidence is of little probative value, however. First: the question framed has a “leading” quality, not well suited to eliciting an uninfluenced reaction from the persons questioned The survey establishes no more than that the names are similar, a factor as to which there can be little genuine dispute in any event, and that portions of the general public will make the reasonable assumption, that, in the absence of any other information, two companies with similar names are likely to have a business connection.¹⁵¹

The issue of business relationships is a tricky one. After all, we *do* want to know whether potential customers falsely believe that there is a connection between, for instance, a less reputable and a more reputable company or product. Yet, as mentioned earlier, similarities in names can be offset by differences that will eliminate confusion. Consumers would also probably be more aware in the marketplace that brand names both containing a generic word (such as here the word “beneficial”) are not necessarily linked. The *Beneficial Corp.* court’s refusal to accept a leading question is intuitively sensible, but it once again shows the dangers in the current judicial approach to trademark surveys that wastes resources and time. In the end, while the evidence here was insufficient for a showing of confusion, it is troublesome that one cannot conclusively state what degree of confusion actually exists in the marketplace, which can potentially have negative repercussions for consumers.

B. *Summary of Second Circuit Doctrine Regarding Survey Contents and Formats*

The majority of acceptable surveys used a visual stimulus that consisted either of the product itself or of an unaltered depiction of the product. Courts also expect surveys to present both the original and the allegedly infringing product, rather than just one of these, while allowing for a variety of settings in which the surveys could take place. It is questionable whether the main exception to the visual stimulus rule, the survey in *McBagel’s*, would be found acceptable if the case were to be decided today. The *McBagel’s* case will hopefully remain an outlier in survey doctrine, but its existence raises the important point that we

¹⁵¹ *Id.* at 450-51; *accord* *Franklin Res., Inc. v. Franklin Credit Mgmt. Corp.*, 988 F. Supp. 322, 335-36 (S.D.N.Y. 1997) (holding that a finance company that used the name “Franklin Credit” did not infringe upon the trademark of a mutual fund company that used the name “Franklin”).

cannot be certain of how an identical case would be decided until courts set a clearer standard for what constitutes an acceptable stimulus. The case has neither been overruled nor has there been any statement on the part of courts that surveys should, for instance, presumptively include as much visual, auditory, and other information as possible to establish confusion.

The Second Circuit has not only failed to state a presumptive standard for survey stimuli but also compounded the problem by failing to state such a standard for control stimuli or the presence of particular environmental cues. Nonetheless, there have been surveys that were struck down for lack of a proper control stimulus or a biasing set of environmental cues. The same is true for the form of survey questions: courts have condemned leading questions, but have not always explained what makes a question leading.

The legal situation of trademark surveys thus rests on an uneasy foundation. There are some indications of measures that researchers can take to make a survey more likely to succeed, such as using the actual product as a stimulus. No survey, however, is guaranteed to succeed, regardless of how carefully prepared or how meticulously it attempts to reproduce those surveys that have been deemed acceptable in the Second Circuit. At the same time, conducting a survey is not a zero-cost game. Not only do experts, subjects, materials, and localities cost a great deal of money,¹⁵² but there is also often no more than one chance for a litigant to conduct a proper survey. Failure to do so can mean losing the entire case. Nonetheless, there is little reason to take a defeatist perspective on the current situation. Clearly, trademark infringement is an area with fuzzy boundaries and no test may perfectly capture “true” infringement. At the same time, courts could take both procedural and substantive corrective measures to remedy the present situation and thus provide the needed guidance to conduct proper surveys and serve the interests of

¹⁵² One figure suggests that these surveys can cost from \$30,000 to \$150,000. See William E. Gallagher & Ronald C. Goodstein, *Inference Versus Speculation in Trademark Infringement Litigation: Abandoning the Fiction of the Vulcan Mind Meld*, 94 TRADEMARK REP. 1229, 1241 n.30 (2004) (citing Keith M. Stolte, *Remedying Judicial Limitations on Trademark Surveys: Monetary Relief Should Not Require Proof of Actual Confusion*, 75 DENV. U. L. REV. 229, 249-50 (1997)). Gallagher and Goodstein discuss whether these high costs should affect judicial requirements of such surveys and conclude:

Regardless of the position with which one may sympathize, it does not advance the inquiry of whether there is a likelihood of confusion in any particular case. It is a matter of policy. If court decisions move the state of the law toward a greater demand for empirical or scientific evidence in cases where the circumstantial evidence is not strong, any unfairness to trademark owners with limited litigation resources is a matter for legislation, not limitations on advancement of decisional law.

Id.

justice.

IV. A NEW MODEL FOR TRADEMARK INFRINGEMENT SURVEYS

As demonstrated, the status of the legal standards surrounding trademark surveys is quite shaky. This Part of the article will argue that a number of steps could improve the existing situation dramatically. First, the current procedural manner in which courts handle trademark surveys should be changed either by having the court itself (through a legislative amendment to the Lanham Act) or the PTO, which currently handles trademark registration, pre-clear surveys before they are conducted. Second, courts need to provide substantive guidance for their expectations, which should include survey rules that would need to be rebutted with specific evidence. An alternative to the latter idea would consist of Congress amending the Lanham Act to include stricter substantive requirements for trademark litigation, but, as explained below,¹⁵³ courts are probably in a better position to implement changes directly.

A. *The Procedural Handling of Trademark Surveys*

Currently, the typical procedure for cases involving trademark surveys roughly works as follows: the plaintiff sues the defendant for a violation of the plaintiff's trademark. The plaintiff hires a researcher or company specialized in conducting trademark-related surveys.¹⁵⁴ Sometimes the defendant himself also commissions a survey to actively disprove the claims of consumer confusion. If the legal case, rather than settling early on, actually goes to trial, a "battle of the experts" tends to ensue. Each side uses experts to defend its own survey while attacking any surveys the other side may have produced. The court must then decide which survey will be admissible, if any. By this point, any mistakes in survey design can be fatal and may not be corrected. Such mistakes can stem from surveyors having attempted to bias subjects a certain way, but errors can also arise out of benign

¹⁵³ See *infra* Part IV.B.

¹⁵⁴ A number of companies now engage in the lucrative business of conducting trademark surveys. See, e.g., *Trademark Infringement Litigation Services*, APPLIED MARKETING SCIENCE, INC., <http://www.ams-inc.com/litigation/trademark.htm> (last visited Oct. 23, 2006) (offering services to establish trademark or trade dress infringement). The company claims:

AMS understands the legal standards for survey research in litigation. We know how to develop questions that are not leading or biased, how to derive valid measurements of confusion, and how to develop stimuli that are truly representative of actual purchase or post-purchase situations. Our 15 years of interviewing customers in dozens of industries have helped us understand the customer mindset, enabling us to "crawl into the minds of consumers."

Id.

motives, such as confusion about existing court standards. One could imagine an alternative system in which the court is involved with the surveys at an earlier stage. Survey scholar Phyllis Welter has suggested that perhaps courts should appoint a third, neutral expert instead of, or in addition to, the parties' experts.¹⁵⁵ Welter quotes Judge Richard Posner, who first proposed a similar idea, to explain the reasons that require such a solution:

Many experts are willing for a generous (and sometimes for a modest) fee to bend their science in the direction from which their fee is coming. The constraints that the market in consultant services for lawyers places on this sort of behavior are weak, as shown by the fact that both experts in this case were hired and, we have no doubt, generously remunerated even though both have been criticized in previous judicial opinions. The judicial constraints on tendentious expert testimony are inherently weak because judges (and even more so juries, though that is not an issue in a trademark case)¹⁵⁶ lack training or experience in the relevant fields of expert knowledge. But that is the system we have.¹⁵⁷

As mentioned, however, it is unclear that potential surveyor bias is the only or even the main problem.¹⁵⁸ There are no guarantees

¹⁵⁵ Phyllis J. Welter, *A Call to Improve Trademark Survey Evidence*, 85 TRADEMARK REP. 205 (1995).

¹⁵⁶ For a brief discussion of the conditions for a jury trial and related trends, which Posner mentions here in passing, see *supra* note 45. The Fourth Circuit has previously stated that juries are particularly competent to resolve questions of likelihood of confusion: "First, the jury, which represents a cross-section of consumers, is well-suited to evaluating whether an 'ordinary consumer' would be likely to be confused. Second, the likelihood of consumer confusion is an 'inherently factual' issue that depends on the unique facts and circumstances of each case." 5 MCCARTHY, *supra* note 10, § 32:130 (citing *Anheuser-Busch, Inc. v. L & L Wings, Inc.*, 962 F.2d 316, 318 (4th Cir. 1992) (citation omitted)).

¹⁵⁷ Welter, *supra* note 155, at 205 (citing *Indianapolis Colts, Inc. v. Metro. Balt. Football Club Ltd.*, 34 F.3d 410, 415 (1994)). Welter also states:

Those clients, the parties, while traditionally interested in saving time and money, are also interested in reducing the risk of a losing position. A mediated survey, done by a neutral court-appointed survey expert, should reduce the overall cost of litigation, with less discovery and shared costs, and should also reduce the risk of a "losing" position by providing an objective evaluation of the market place.

Id. at 209-10.

¹⁵⁸ It is difficult to estimate how many of the problems with survey evidence were caused by researcher bias as opposed to other issues. Some judges have, however, expressed specific concern about the issue of bias:

A universe that excluded whites entirely is one which reflects neither the market for rap music generally nor for plaintiffs' music specifically. An all-black universe reflects the readership of the magazine, not the listeners of rap; in effect, the answers rendered by that universe are responding to the question of whether consumers of the magazine are likely to be confused. Only if whites are included will the answers reflect whether consumers of rap performers are likely to be confused. It is thus not a proper universe. Given the extensive testimony about the powerful name recognition enjoyed by ESSENCE Magazine amongst blacks, the exclusion of whites was likely to bias the results in favor of the magazine.

that a neutral expert would necessarily produce a survey that conforms to what courts want. While courts may, for better or for worse, be more inclined to trust the judgment of a neutral expert, his survey may still not fulfill certain expectations. Furthermore, it has been noted that “judges as well as lawyers want trials to be predictable and controlled,”¹⁵⁹ a goal that the appointment of a neutral expert would undermine. One scholar explains:

A court-appointed expert . . . is nobody’s responsibility. Preparation by the parties’ attorneys will not have the usual effect. They have no control over an appointed expert and are likely to be suspicious of the entire enterprise, if not overtly hostile, while the expert herself has no commitment to any one party and will probably be reluctant to compromise her neutrality by working too closely with either side. Worse, at least some authorities believe that *ex parte* communications between attorneys and court-appointed experts are improper. This would mean that any pre-trial preparation by the attorneys must take place (if at all) in the presence of the opposition—a procedure American lawyers will shun.¹⁶⁰

The scholar concludes that having a neutral expert would produce a transformation of the adversarial system that would “threaten the prerogatives of the trial attorneys”¹⁶¹ and produce wide-spread opposition on the part of the bar¹⁶² as well as judges, who are often former trial attorneys themselves and consequently share some similar sentiments.¹⁶³ Thus, while a system with neutral experts presents certain advantages, it is unlikely to receive the necessary amount of support from the legal community in any case.¹⁶⁴

Some scholars have even suggested that there are distinct advantages to having surveys specifically conducted by the plaintiff and defendant. In their article about this exact issue,¹⁶⁵ Ruth

Hutchinson v. Essence Commc’ns, Inc., 769 F. Supp. 541, 564 (S.D.N.Y. 1991).

¹⁵⁹ Samuel R. Gross, *Expert Evidence*, 1991 WIS. L. REV. 1113, 1202 (1991).

¹⁶⁰ *Id.* at 1203.

¹⁶¹ *Id.* at 1205.

¹⁶² *Id.* at 1197-98.

¹⁶³ *Id.* at 1198.

¹⁶⁴ Gross also mentions some problems that are specific to the court appointing psychologists or psychiatrists, i.e., that “these are highly controversial areas for expert evidence in any form, probably the most controversial subjects on which expert testimony is routinely allowed,” that “psychiatric and psychological expertise is . . . different from other types of expertise and . . . is inherently partisan,” and that “in some jurisdictions in which court-appointed psychiatrists and psychologists are used comparatively frequently, they are typically chosen for reasons of convenience and economy rather than quality.” *Id.* at 1195. While Gross’s discussion primarily appears to focus on the criminal law context, these factors could cause at least partial problems in the trademark survey context as well.

¹⁶⁵ Ruth M. Corbin & Arthur Renaud, *What’s Good for the Goose Is Bad for the Gander: Why Confusion Surveys for Plaintiff and Defendant Should Be Different*, 16 INTELL. PROP. J. 177 (2003).

Corbin and Arthur Renaud name a number of reasons defending this status quo by focusing on Canadian case law but using arguments equally applicable in the American context. Corbin and Renaud note that the two parties must meet opposite evidentiary burdens—plaintiffs must demonstrate likelihood of confusion while defendants must demonstrate that there is none—which calls for different survey formats for scientific and statistical reasons.¹⁶⁶ In the end, “there are legitimate situations in which each side’s evidence is not contradictory, but differently directed. It is then up to the Court to seek to understand what each side’s evidence demonstrates, and how each bears on the issues at hand.”¹⁶⁷ Corbin and Renaud also discredit the idea that surveys are some sort of “junk science” that demonstrate whatever their creators wish them to.¹⁶⁸ Rather, “the field of survey research incorporates all the essential structural techniques of other scientific expert evidence, including rigorous hypothesis testing, experimental design, control conditions, and statistical inference,”¹⁶⁹ and the scholars view it as an advantage that “lawyers will have the opportunity to educate the court on what constitutes proper statistical inference, as opposed to pseudo-science, in defending their own survey strategy and criticizing the opposition.”¹⁷⁰

Despite these criticisms of neutral experts and defenses of the status quo, an interesting aspect of Phyllis Welter’s proposal¹⁷¹ is the idea of allowing judges to intervene before any survey is conducted. One way to implement this idea would be to require

¹⁶⁶ *Id.* at 182-84.

¹⁶⁷ *Id.* at 184.

¹⁶⁸ *Id.* at 178.

¹⁶⁹ *Id.* at 180.

¹⁷⁰ *Id.* at 194. The piece also notes that judges in Canada, like in the United States, have rejected trademark confusion claims altogether for failure to present trademark evidence. *Id.* at 179. Another scholar states:

In a number of other countries, mostly in Europe for the time being, the trademark bar is increasingly resorting to surveys. In Germany and the United Kingdom, surveys are by now quite common, although some British judges still regard them with a healthy dose of skepticism, if not hostility. In Sweden, the Patent Office recently withdrew an objection of non-registrability following the submission of survey evidence. In Austria, survey evidence is accepted both by the Patent Office and the courts. The Finnish Supreme Court two years ago decided the ADIDAS “three stripe” case on the basis of survey evidence.

Peter Weiss, *The Use of Survey Evidence in Trademark Litigation: Science, Art or Confidence Game?*, 80 TRADEMARK REP. 71, 73 (1990) (footnotes omitted). Roland Knaak explains that in Germany, judges can issue a court order demanding a survey if they feel that *their own* experiences as consumers are insufficient to reach a result. Roland Knaak, *The International Development of Survey Practices in Trademark and Unfair Competition Law*, 21 INT’L REV. INTELL. PROP. & COMPETITION L. 327, 331 (1990). Knaak further predicts that “it appears unlikely that such [likelihood of confusion] surveys will be frequently applied in the near future” due to the “complexity of the likelihood of confusion in its legal sense and in its actual types.” *Id.* at 343.

¹⁷¹ See *supra* note 155 and accompanying text.

parties to submit survey proposals to the judge before conducting the actual surveys. Without an argument between the parties, judges could thus establish in advance if the survey was acceptable *on its face*. This would require the court to look at the survey questions, proposed setting, survey universe and the like, and to make a preliminary decision about whether a particular survey framework would ever be deemed acceptable. The reason this would only constitute facial acceptability is that there could still be methodological errors in implementation that would make the survey unreliable.

It appears, however, that most of the surveys that were rejected by the Second Circuit would have already been rejected at this pre-trial stage, which would have given the relevant parties the opportunity to improve the design until it directly met the court's standards. This includes, for instance, the survey in the "Charbucks" case previously mentioned.¹⁷² Considering the fact-specific nature of trademark infringement law, judicial pre-clearance would produce an efficient way to ensure qualitatively high research that would presumptively be approved by the court before parties spend high sums of money on surveys. This idea presents certain key advantages over Welter's model: it still allows the parties to choose the experts they deem the most suitable and ensures that most methodological mistakes will be caught at an early stage. One potential criticism of pre-clearance of surveys is that it could force judges to spend a lot of time examining several versions of a survey. However, it appears that parties have little incentive to "annoy" a judge with too many versions and will try to minimize their demands on him. This approach would also be coupled with substantive guidelines¹⁷³ to reduce the strain on the judicial system and encourage an efficient allocation of resources.

Procedurally, and resonating with Welter's thought, one could establish these guidelines with the help of numerous, neutral trademark survey experts that incorporate some very general bright line rules to create a floor below which no survey could presumptively fall if it is to be declared facially acceptable. It is understood that social science will never become as exact as other sciences, but a stated rule reasonably based on available, albeit slightly inexact knowledge, may prove better than no rule at all. Experts could still present empirical evidence of why the subject they studied is so extraordinarily different as to merit the rebuttal of the court's presumptions. While the empirical feasibility and precise outline of procedural changes remains to be

¹⁷² See *supra* Part I.

¹⁷³ See *infra* Part IV.B.

explored in further detail by future scholars, it is clear that it should become one of the most important contexts in which to consider improvement of the judicial treatment of surveys.

It also bears emphasis that certain elements of a survey, such as survey stimuli, are particularly easy to subject to pre-trial judicial scrutiny.¹⁷⁴ Researchers can submit to the court the exact materials they are planning to use and courts can acquiesce to the chosen stimulus, reject the selection, or ask for its modification in a particular way. Hearings on these issues would be conducted outside the presence of the jury to prevent prejudice. While the court at this point only establishes facial admissibility and will determine final admissibility of the survey later, the jury will still need to evaluate the weight it chooses to lend to the survey. If a survey does make its way to the jury after being deemed facially acceptable, and if the court explains the procedure that permitted this, juries will understand that the survey met a basic set of criteria that correspond to the current state of social science knowledge. Juries in different trademark infringement cases would thus consider surveys that contain fairly homogenous basic elements, which would ensure more equal treatment across litigations. Knowing that they are examining a survey whose questions, stimuli, and other founding blocks can be trusted would allow juries to focus their attention on whether a facially acceptable survey was also executed properly and, if so, how its results fit with the totality of the evidence.

Alternatively, Congress could amend the Lanham Act to allow the PTO to pre-clear surveys in an agency proceeding. The PTO already possesses significant expertise in making decisions related to questions of infringement because of its decision-making role in the trademark registration process. When an individual seeks to register a trademark, other parties who believe the registration would damage them have the right to file an opposition to registration with the PTO, after which an agency determination decides the question.¹⁷⁵ Additionally, parties can file an application with the PTO for cancellation of an already registered mark.¹⁷⁶ Further investigation needs to be performed to quantify the potential cost of PTO pre-clearance and the degree to which the agency's operation would have to be modified before it could fulfill this task. Until this data is available, judicial pre-clearance

¹⁷⁴ In addition, survey elements such as control stimuli, question formulations, definition of the survey universe, and others are also easy to subject to pre-trial judicial scrutiny.

¹⁷⁵ See 15 U.S.C. § 1063 (2006).

¹⁷⁶ See *id.* § 1064.

represents the most straightforward procedural change: while its exact costs would still require investigation, the judicial system would demand less extensive change to perform such a function.¹⁷⁷

B. *Substantive Guidelines for Content-Related Issues*

Would it not be useful if judges could agree on certain minimal presumptive conditions that every survey must fulfill before being admitted? Would it not be even better if courts could elaborate and explain what these conditions are rather than make experts rely on case-by-case judgments? The answer to these wishes may in fact lie in the realm of the possible. The *Toys “R” Us*¹⁷⁸ and similar tests tried to do precisely that, but they left too many open questions, which resulted in the uneven treatment of surveys across cases and an endangerment of reliance values in the legal system. A number of scholars have debated in what conditions *rules*, which create *ex ante* requirements, are more efficient than *standards*, which establish *ex post* requirements.¹⁷⁹ The most important factor in determining the efficiency of each of these two methods for a given situation is the frequency with which a certain rule or standard will have to be applied—rules tend to be more efficient in situations in which a great number of similar cases exist.¹⁸⁰ This is clearly true of trademark infringement surveys: they are a necessary component of proving trademark infringement in a great number of lawsuits and they generally resemble each other in many ways.¹⁸¹ Furthermore, establishing the *ex ante* rules proposed in this article would not offer additional opportunities to “game” the judicial system through creative lawyering¹⁸² because following the rules would

¹⁷⁷ Professor Douglas Lichtman raised the question to me whether, in a model that involves PTO pre-clearance of surveys, the PTO would have a role in helping the parties to choose and prepare their experts. I would opt against giving the PTO this role under the rationale that this would be 1) a presumably expensive addition to my proposal; 2) a profound change to today’s PTO (whereas examining questions directly tied to infringement relates to its current function in choosing what marks to register or cancel); and 3) unnecessary; if proper guidelines for survey constructions existed, the free market would likely allow for appropriate expert selection and experts would need no further guidance than lawyers themselves can provide in a consistent legal system.

¹⁷⁸ See *supra* Subpart III.A.1.

¹⁷⁹ See, e.g., Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557 (1992).

¹⁸⁰ See *id.* at 585-86. Kaplow states specifically:

Whether the ideal time to acquire and disseminate information is *ex ante* or *ex post* depends, most importantly, on the frequency with which the information will be used. The savings from a single *ex ante* investigation will be great when the use of the results will be frequent, but will be negligible when the use of the results will be unlikely.

Id. at 585.

¹⁸¹ The resemblance amongst surveys is part of the larger argument in this article: there is a social science consensus as to the construction of some key building blocks of surveys.

¹⁸² One of the main areas in which creative lawyering is a problem is tax law. As one

not provide any guarantee of success in the courts. Rather, the rules would only establish a floor below which the quality of a survey could not fall to be considered in any further proceedings.

This Part proposes several concrete requirements that courts should set in the form of substantive rules and ideally combine with the procedural changes outlined in Part IV.A of this article. Theoretically, there is the possibility that Congress could delineate some of these standards in an amendment to the Lanham Act under the rationale that members of Congress could pull together information from experts and develop a uniform framework for trademark litigation. This idea, however, presents several difficulties. First, it would introduce an unprecedented type of detail into the Lanham Act. Second, at times a certain level of flexibility or exceptions is reasonable, and fixating guidelines in the statutory framework itself rather than allowing judicial presumptions could introduce too much rigidity. Third, progress in social science research methods may require adjustments to any rules in the future, and requiring Congress to amend the Lanham Act whenever that occurs could prove unwieldy and impractical. Thus, courts are probably in a better position to implement substantive guidelines.

1. The Proper Stimulus

Courts should presumptively require surveys to include either the product itself or an accurate *visual* depiction of the product as the stimulus presented to subjects. Most of the surveys that the Second Circuit has found acceptable already used a visual depiction of the product, so this in no way represents a major change to survey requirements but is rather a clarification in doctrine.¹⁸³ The decision of whether to use the actual product *or* another depiction should depend on whether the products in question possess other characteristics such as texture or smell that cannot be conveyed by a mere image.¹⁸⁴ Because consumers take in the totality of a product in the marketplace, an image alone would be insufficient if there is any doubt that texture or smell would make a difference in the totality of factors that point toward

scholar explains: "Because the system focuses on transactions rather than on income, and because those transactions themselves are not always defined in a way that correlates with their underlying purpose, the system is easily subject to manipulation." Joshua D. Rosenberg, *Tax Avoidance and Income Measurement*, 87 MICH. L. REV. 365, 444 (1988).

¹⁸³ It is understood that courts would presumptively continue to require that both the original and the allegedly infringing product be presented as part of the survey.

¹⁸⁴ In an Internet survey, for instance, sound could easily be incorporated into the study as well, while these other attributes could not. Policy reasons may also exist to not show survey subjects actual products of a threatening nature, such as guns, and thus accurate images may prove more appropriate in those cases.

or away from the existence of a trademark violation.

This approach would allow researchers to know definitively what stimuli will be acceptable on their face in trademark infringement cases, and, more importantly, what stimuli will not. Special problems would still exist for other forms of confusion, such as initial interest confusion, because that type of confusion may occur when subjects are still at a distance from a product and may only see some of its elements. For instance, a soda bottle with black liquid in it and a red-and-white label may mistakenly attract a consumer because he believes it to be Coca Cola, but when the consumer steps closer to the product, it may turn out to be a product by another brand. This may hurt the consumer in several ways, such as by wasting his time.¹⁸⁵ The fact that only certain elements are visible in the marketplace should, however, not prevent researchers from using the actual product with all markings in a survey. In the soda bottle example, researchers would simply have the choice to use the actual product as a stimulus and ask consumers to stand a few feet away from it the way they would at a supermarket. Disallowing modified products in trademark surveys would thus not stand in the way of uncovering initial interest confusion in this or other cases through similar, specific measures that researchers are still free to take.

If researchers choose to use an image of a product, they should be required to display all the facets of the product that a consumer would encounter in the marketplace. Ideally, in Internet surveys, researchers could easily use modern “virtual tour” technology that allows the display of an object and manipulation by mouse to view all three dimensions. Because in-person studies can understandably present considerable expenses, the question of confusion could thus be handled properly on the Internet for many products.

¹⁸⁵ See Michael Grynberg, *The Road Not Taken: Initial Interest Confusion, Consumer Search Costs, and the Challenge of the Internet*, 28 SEATTLE U. L. REV. 97, 109 (2004). Grynberg states:

[I]nitial interest confusion is not a guided missile precisely targeting only those consumers who will decide that the alleged infringer’s product is superior to that of the trademark holder’s. It also affects consumers who wish to purchase the trademarked brand and must expend extra effort to find it. Their confusion, though ultimately dispelled, increases search costs even when they ultimately purchase from the trademark owner. Also, many of the diverted consumers would have preferred the initial target of their search, but settle for the infringer’s product because finding the trademark holder would be too much trouble. It is these increased consumer search costs that justify judicial attention to initial interest confusion.

Id. (footnote omitted).

2. The Proper Control Stimulus

Just as judges can approve of survey stimuli on their face prior to trial, they can do the same for control stimuli. The control stimulus should be presented in the same medium as the survey stimulus, so that a survey stimulus in the form of a physical product would be matched with another physical product, an image with another image, and so forth. One possible judicial presumption could be that an actual, existing control product or image should be used such as to counter the possibility to manipulate the control stimulus. At the same time, this may not be feasible for every market if such a product is not available, or if only one is available that would produce no confusion at all.¹⁸⁶ As stated earlier, “if in a particular test, a control stimulus is so different that no respondent would ever guess it, even when it is the only stimulus shown, then such a control cannot measure whether people are guessing with response to the stimulus of actual interest”¹⁸⁷ A potential compromise could be to set the judicial presumption at demanding an existing product to be used as the control stimulus, but only require an intermediate standard for rebutting this presumption as opposed to the high standard that would be required to rebut the presumptions discussed in the survey stimulus context.¹⁸⁸

3. The Proper Questions

Courts need to define the types of questions that will be acceptable in the trademark context. Fortunately, the procedural component of the proposal in this article would correct any problems in question formulation at the pre-trial stage. The most important lesson to learn from the Second Circuit’s analysis of trademark-related cases is that questions need to be precise¹⁸⁹ and

¹⁸⁶ See, e.g., *Empresa Cubana Del Tabaco v. Culbro Corp.*, No. 97 Civ. 8399, 2004 WL 602295, at *26 (S.D.N.Y. Mar. 26, 2004) *rev’d in part*, 399 F.3d 462 (2d Cir. 2005), *cert. denied*, 126 S. Ct. 2887 (2006) (“It is difficult to determine . . . exactly how a control could have been helpfully employed in the . . . survey. Any non-COHIBA word trademark would not serve as an effective control, while a fictional third COHIBA cigar would clearly infringe on the senior mark.”).

¹⁸⁷ See Rappeport, *supra* note 62, at 987. This means that if someone were, for instance, to claim that Pepsi Cola bottles and their content infringe on the equivalent Coca-Cola products, the actual survey stimulus would be a full Pepsi Cola bottle and the control stimulus would have to consist of a product at least somewhat similar to the Pepsi bottle rather than of an orange juice carton; after all, there is little evidentiary value in the fact that more people would confuse a Pepsi bottle with a Coca-Cola bottle than would confuse an orange juice carton with a Coca-Cola bottle.

¹⁸⁸ See *supra* Part IV.B.1.

¹⁸⁹ See, e.g., *Am. Footwear Corp. v. Gen. Footwear Co.*, 609 F.2d 655, 660 (2d Cir. 1979) (rejecting the survey question “With whom or what do you associate a product labelled Bionic?” which should have been formulated “With whom or what do you associate a ‘Bionic’ boot?”); see also Artemio Rivera, *Testing the Admissibility of Trademark Surveys After*

non-leading.¹⁹⁰ Accordingly, it appears sensible to set the following judicial presumptions: first, questions must contain the exact name of the potentially infringing product or otherwise clearly identify the survey stimulus.¹⁹¹ Second, the question may not suggest on its own the brand or other product with which one may associate the survey stimulus. A proper question would thus ask “Who do you think makes this bottle?” rather than “Is there any chance this bottle is made by Coca-Cola?”¹⁹² Third, questions may not use any adjectives to qualify a term and thus attempt to hint at an answer, such as “What large company do you think makes this bottle?” Ideally, in fact, the question should not even hint at whether a product is manufactured by an individual versus a company. A good question would ask “Who makes this bottle?” rather than “What company makes this bottle?” There are no clear disadvantages to the question types proposed here as long as they are stated in precise, understandable terms. Fourth, surveys should ask follow-up questions in which it could become apparent what caused any existing confusion.¹⁹³ These should take the form of “What makes you say that?” or “Why is that your answer?”¹⁹⁴

Daubert, 84 J. PAT. & TRADEMARK OFF. SOC'Y 661, 680-81 (2002) (“Survey questions must be clear and use wording that [is] easy to understand by the average person in the target population. . . . Since factors such as age, sub cultural group, and region can affect the meaning of a word, researchers should use simple language and avoid wording that can have multiple interpretations.”).

¹⁹⁰ See, e.g., *Beneficial Corp. v. Beneficial Capital Corp.*, 529 F. Supp. 445, 450 (S.D.N.Y. 1982) (rejecting the question “Do you think that there may or may not be a business connection between Beneficial Capital Corp. and the Beneficial Finance System Companies?”).

¹⁹¹ Another clear way of identifying the survey stimulus would be by stating, for instance, “Who do you think made the product labeled A?”

¹⁹² Rivera distinguishes certain situations, however, in which questions should not be entirely open-ended and cautions how experts must vary the order of questions if they use close-ended ones. Rivera, *supra* note 189, at 683-85.

¹⁹³ Shari Seidman Diamond has in fact criticized the failure of surveys to ask follow-up questions in the mail survey context by stating that “[o]pen-ended questions are generally of limited value on a mail survey because they depend entirely on the respondent to answer fully and do not provide the opportunity to probe or clarify unclear answers.” Shari Seidman Diamond, *Reference Guide on Survey Research*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 229, 263-64 (2d ed. 2000). Asking the follow-up questions proposed in this article would at least partly alleviate this problem. Another step (discussed *infra* Part IV.B.4) would be to suggest further post-study follow-up for certain studies.

¹⁹⁴ A few more issues surround the requirements for proper questions. Importantly, there has been disagreement about the best way to discourage people from guessing the answers to survey questions. For an explanation of why introducing an admonition against guessing or an “I don’t know” option may not be helpful, see Rappeport, *supra* note 62, at 984 (footnote omitted):

It is frequently claimed that giving a respondent an initial instruction “not to guess,” or “not to be afraid to say they don’t know” will, in fact, reduce the tendency to guess. . . . [I]n the real world, such an admonition generally does not have any discernible effect on respondents. This absence of an effect is an outgrowth of the underlying reason for a respondent’s guessing, namely, that respondents assume that there are answers to the questions posed. Given this assumption, an admonition “not to guess” is as likely to confirm for some respondents that there is a “right” answer that they should be smart enough to

4. The Proper Environment

The survey should be conducted in the same setting or medium in which the sale of the potentially infringing product would occur. Demonstrating confusion as part of an Internet study would thus be acceptable for products sold over the Internet or equivalent contexts such as mail-order catalogues. The Internet would, however, not be acceptable for products only sold, for instance, in supermarkets, because of the large variation of environmental cues in the two media. Yet, this presumption could be overcome if data could demonstrate that an in-person study is not feasible and if the Internet study yielded large amounts of confusion. Phone studies would be eliminated, with the possible exception of surveys for products available through direct telemarketing. Any settings that do not involve in-person contact would have to include a mechanism to ask further follow-up questions if the answers are unclear after the first set of follow-up questions already in the original study. For instance, one could mandate follow-up questions for any research subjects that did not explicitly answer according to the directions in the study.¹⁹⁵

5. Discussion

The procedural and substantive components of the proposal represent two discrete sets of ideas that courts could implement either separately or together in an attempt to unify the treatment of trademark surveys, decrease costs, and increase people's ability to rely on a body of clear law. While this article recommends using both components together, there is nothing illogical about choosing one or the other. There are probably many variations and alternatives to the suggestions made here. The main point is, however, that trademark law would benefit from clarification and that courts should study how this goal can be achieved, whether through the proposal in this article or by other means.

It is always problematic when an area of law lacks precise

know, and thus to encourage guessing, as it is to discourage guessing among other respondents. . . . This general admonition against guessing sometimes is manifested as an insistence upon the value of including as an answer category to individual questions: "I don't know." Including an "I don't know" option is . . . less likely to encourage guessing and thus may be a harmless practice, although not of real value.

Id. But see Jacob Jacoby, *A Critique of Rappeport's "Litigation Surveys—Social 'Science' As Evidence*, 92 TRADEMARK REP. 1480, 1482-87 (2002) (disagreeing with Rappeport). Another alternative is the introduction of so-called "full-filter questions" in which subjects are first asked if they know about the substantive subject matter of the survey and questioning only continues if they express that they do. See Rivera, *supra* note 189, at 681-82. Rivera also explores when control questions are appropriate. *Id.* at 686-88.

¹⁹⁵ While these proposals do not specifically discuss the topic of trademark dilution, they are all applicable to dilution surveys as well.

definitions, particularly when the related cases involve great amounts of money in damages and other costs. While the world of potential trademark violations may never allow for black-and-white decisions because of the necessity of examining the mental state of consumers, the possibility to set certain boundaries does exist. This should happen by making explicit the guidelines that courts seem to have employed in most cases and clarifying ambiguous areas of the law. The larger goal here is to identify and prevent true trademark infringement while disallowing the monopolization of an entire industry through over-protection. The only way to advance toward that goal is to standardize the use of trademark survey evidence, such that laws, rather than men, can dominate the field. It has been stated in the criminal law context: "Truth, though unattainable in all of its labyrinthic extremities, must always be the judicial goal. It is the purpose undergirding our rules of evidence."¹⁹⁶ Courts should strive to follow this principle to the best of their abilities in the civil context as well and define their rules of operation accordingly.

V. CONCLUSION

Trademark infringement surveys have had a tumultuous history: first deemed inadmissible, then admitted and later transformed by the *Daubert* standard, they are now one of the touchstones of trademark infringement litigation. Courts' treatment of surveys, however, remains riddled with inconsistencies, and creates insecurity for litigators as well as survey constructors. This article exemplifies the current status of trademark surveys by extensively discussing the way in which courts within the Second Circuit have handled surveys generally, and their contents specifically. To remedy the lack of guidance that litigants receive in the areas of survey stimuli, questions, and environments, this article presents a model of substantive guidelines that courts should establish to determine survey admissibility. Ideally, courts would implement a procedure through which judges or the PTO could pre-screen whether surveys meet these guidelines before the surveys' actual execution. This would allow judges to declare whether particular surveys are admissible on their face, which would cut the costs involved in conducting the most unreliable types of surveys and avoid many surprises regarding admissibility during trial.

The model suggested here would benefit from further research into the details of both its substantive and procedural

¹⁹⁶ State v. Sheppard, 484 A.2d 1330, 1343 (N.J. Sup. Ct. 1984).

components. This article points to the deficiencies in the current system but remains open to different possibilities as to the exact degree to which judges or the PTO should scrutinize surveys before declaring them admissible on their face, the procedure through which litigants should submit survey proposals and the number of times they should be able to do so, as well as several other questions. While each of these issues warrants examination and empirical research regarding the efficiency and feasibility of various solutions, this article aims to open the dialogue about an area whose inconsistencies much of the scholarly literature has overlooked for too long.