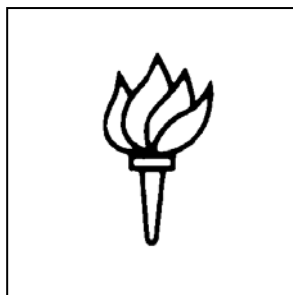


# NEW YORK UNIVERSITY SCHOOL OF LAW

NYU Center for Law, Economics and Organization

**LAW & ECONOMICS RESEARCH PAPER SERIES**

**WORKING PAPER NO. 18-06**



Testing for Trademark Dilution in Court and the Lab

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

# Testing for Trademark Dilution in Court and the Lab

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

## Abstract

Federal courts are currently split, even within particular districts, on the basic question of what a plaintiff must show to establish that a defendant's conduct constitutes trademark dilution by blurring. Federal trademark law defines "dilution by blurring" as "association arising from the similarity between a mark or trade name and a famous mark that impairs the distinctiveness of the famous mark." In construing this statutory language, a majority of courts have held that to establish blurring, a plaintiff need only show that consumers associate the defendant's mark with the plaintiff's famous mark. These courts appear to assume that to the extent that there is consumer association, this association alone will "impair the distinctiveness" of the famous mark. A minority of courts have held that the plaintiff must show *both* consumer association *and* that this consumer association "impairs the distinctiveness" of the famous mark. In this article, we make two contributions to the current debate over what must be shown to establish dilution by blurring. First, we report the results of a set of experiments that reveal that the majority approach is fundamentally deficient. These experiments demonstrate that even when consumers associate a junior mark with a famous senior mark, this association does not necessarily result in any impairment of the ability of the senior mark to identify its source and associations. Second, we evaluate the current state of the art in trademark dilution survey methodology: response time surveys. These surveys purportedly show both consumer association and impairment. Through a set of experiments, we demonstrate that these surveys currently use the wrong control and are invalid. In light of our findings, we reflect more generally on the question of whether dilution by blurring ever occurs and on how the blurring cause of action may be reconfigured better to comport with courts' intuitions about the true nature of the harm that the cause of action seeks to address.

---

\* Barton Beebe, John M. Demarais Professor of Intellectual Property Law, New York University School of Law; Roy Germano, Research Scholar, New York University School of Law; Christopher Jon Sprigman, Professor of Law, New York University School of Law; Joel H. Steckel, Professor of Marketing and Vice Dean for Doctoral Education, NYU Stern School of Business. The authors thank Jamie Boyle, Chris Buccafusco, Alex Cadmus, Shari Diamond, Michael Frakes, Jeanne Fromer, Scott Hemphill and Stephan Tontrup and participants in workshops at the New York University School of Law, the Duke University School of Law, the St. John's University School of Law, and the 2017 Intellectual Property Scholars Conference for helpful comments and conversations. Thanks also to Ari Lipsitz for excellent research assistance.

Introduction.....	3
I. Mere Association and Blurring.....	7
A. Survey Evidence of Mere Association .....	7
1. The <i>Nikepal</i> Survey Approach .....	7
2. Association and Confusion Survey Results.....	9
B. Survey Evidence of Mere Association is Not Evidence of Impairment ..	11
1. Brand Selection Pretest .....	11
2. Study 1: Examining Brand Association Strength.....	12
3. Study 2: Dilution at the Cash Register .....	14
II. Response Time Experiments and Blurring.....	16
A. Previous Response Time Experiments .....	16
1. Morrin and Jacoby.....	17
2. Pullig, Simmons, and Netemyer.....	19
B. Response Time Pilot Studies: Are We Really Observing Dilution? .....	20
1. Pilot Study 1: The Surprise Effect.....	21
2. Pilot Study 2: Controlling for Potential Surprise .....	23
C. Study 3: Response Time and Category Recall When Controlling For Surprise .....	25
1. Response Times.....	25
2. Category Recall .....	28
III. Does Trademark Blurring Ever Occur? .....	32
A. Superdilution? .....	33
B. Dilution and Misappropriation .....	33
IV. Conclusion: Future Directions for Research.....	35
Appendix A: Amazon Mechanical Turk .....	37
Appendix B: Advertisements Used in Studies 1, 2, and 3 .....	39

## Introduction

Trademark dilution is among the most elusive concepts in intellectual property law.<sup>1</sup> In theory, dilution occurs when two companies use very similar or even identical trademarks in a manner that does not confuse consumers as to source. Consumers are instead aware that the two marks refer to two different companies. The problem is that these consumers must “think for a moment”<sup>2</sup> whenever they see one of the marks to determine to which company the mark refers, and this is thought to represent a harm to the prior, senior user of the mark. For example, if a Brooklynite opens in her neighborhood a microbar called “Apple,” it is highly unlikely that patrons will be confused into believing that the global high-technology brand owns the bar and is leveraging its brand into the sale of artisanal cocktails. There is no consumer confusion as to source. But the concept of dilution asserts that the existence of the bar may nevertheless harm the famous global brand. When Brooklynites hear the term “Apple,” they will associate it with two different entities: is it the high-tech company (the senior user of the trademark), or the bar (the junior user)? As the legal theory goes, this dual association causes a “blurring” of the link between “Apple” and the Cupertino company. This “dilution by blurring” is understood somehow to damage the famous brand name by diminishing the immediacy with which consumers identify it with its source and associations.

---

<sup>1</sup> See J. Thomas McCarthy, *Dilution of a Trademark: European and United States Law Compared*, 94 TRADEMARK REP. 1163, 1163 (2004) (“No part of trademark law that I have encountered in my forty years of teaching and practicing IP law has created so much doctrinal puzzlement and judicial incomprehension as the concept of ‘dilution’ as a form of intrusion on a trademark.”).

<sup>2</sup> Richard Posner, *When Is Parody Fair Use?*, 21 J. LEGAL STUD. 67, 75 (1992) (explaining antidilution law on the basis that “a trademark seeks to economize on information costs by providing a compact, memorable and unambiguous identifier of a product or service. The economy is less when, because the trademark has other associations, a person seeing it must think for a moment before recognizing it as the mark of the product or service.”). Judge Posner reasons similarly in explaining the rationale for protection specifically against dilution by blurring:

[T]here is concern that consumer search costs will rise if a trademark becomes associated with a variety of unrelated products. Suppose an upscale restaurant calls itself “Tiffany.” There is little danger that the consuming public will think it’s dealing with a branch of the Tiffany jewelry store if it patronizes this restaurant. But when consumers next see the name “Tiffany” they may think about both the restaurant and the jewelry store, and if so the efficacy of the name as an identifier of the store will be diminished. Consumers will have to think harder—incur as it were a higher imagination cost—to recognize the name as the name of the store. So “blurring” is one form of dilution.

*Ty Inc. v. Perryman*, 306 F.3d 509, 511 (7th Cir. 2002) (citations omitted).

The concept of trademark dilution is as controversial as it is vague. Trademark scholars are overwhelmingly critical of antidilution protection, which appears to grant in-gross rights to owners of qualifying marks<sup>3</sup>—such that Apple (the technology company) could use antidilution law to prevent nearly any subsequent user from adopting the term “apple” as a trademark, even when no consumer confusion plausibly results. Commentators also question whether ostensibly diluting conduct causes any real harm. Christine Haight Farley defies proponents to provide even a single concrete (and not hypothetical) example of a mark that has been significantly damaged because another firm has used that mark in a nonconfusing manner on different goods.<sup>4</sup> Indeed, it is far from obvious how one establishes whether dilution even occurs, not just in a particular litigation, but whether it ever even occurs at all.<sup>5</sup> How would one go about showing that the boîte in Brooklyn has harmed in any way the world-famous brand with which it shares a name? Judges too remain highly suspicious of antidilution law. As one federal appellate judge put it to an antidilution plaintiff at oral argument, “boy you must have some lobby to get a law like that passed.”<sup>6</sup>

Making matters worse, federal courts are currently split on the basic question of what the plaintiff must show to establish that the defendant’s conduct constitutes blurring. Even within particular districts, such as the Southern District of New York, courts have adopted different standards.<sup>7</sup> Federal trademark law states that “‘dilution by blurring’ is association arising from the similarity between a mark or trade name and a famous mark that impairs the distinctiveness of the famous mark.”<sup>8</sup> In construing this language, a majority of courts have held

---

<sup>3</sup> See Bd. of Regents, Univ. of Texas Sys. ex rel. Univ. of Texas at Austin v. KST Elec., Ltd., 550 F. Supp. 2d 657, 674 (W.D. Tex. 2008) (“Dilution causes of action, much more so than infringement and unfair competition laws, tread very close to granting “rights in gross” in a trademark.”).

<sup>4</sup> See Christine Haight Farley, *Why We Are Confused about the Trademark Dilution Law*, 16 FORDHAM INTELL. PROP. MEDIA ENT. L. J. 1175, 1184-85 (2006).

<sup>5</sup> See Shari S. Diamond, *Surveys in Dilution Cases*, in TRADEMARK AND DECEPTIVE ADVERTISING SURVEYS 159, 160 (Shari S. Diamond & Jerre B. Swann eds., 2012) (“[W]e do not have survey methods that can measure when a particular association is likely to lead to impaired distinctiveness of a particular mark. . . . The result is that courts both now and in the foreseeable future will have to continue to struggle, unaided by direct survey evidence, in making predictions about the ultimate question: likelihood of dilution.”). See also J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 24:121 (5th ed. 2017).

<sup>6</sup> Transcript of Oral Argument, *Ringling Bros.-Barnum & Bailey v. Utah Div. of Travel Dev.*, 935 F. Supp. 763 (E.D. Va. 1996).

<sup>7</sup> Compare *Louis Vuitton Malletier, S.A. v. Hyundai Motor Am.*, 2012 WL 1022247 (S.D.N.Y. 2012) (finding proof of association conclusive evidence of dilution), with *Louis Vuitton Malletier, S.A. v. My Other Bag, Inc.*, 156 F.Supp.3d 425, 436 n.4 (S.D.N.Y. 2016), *aff’d*, 2016 WL 7436489 (2d Cir. 2016) (rejecting the *Hyundai* reasoning, stating “association is a necessary, but not sufficient, condition for a finding of dilution by blurring”).

<sup>8</sup> 15 U.S.C. § 1125(c)(2)(B).

that to establish blurring, a plaintiff need only show that consumers associate the defendant's mark with the plaintiff's famous mark.<sup>9</sup> In other words, to win an injunction, Apple of Cupertino would need only show that patrons of the Brooklyn bar associate its name with the Cupertino company (even if there is no source confusion). These courts appear to assume that to the extent that there is association, this association alone will "impair the distinctiveness" of the famous mark.<sup>10</sup> We refer to this in what follows as the *mere association* standard. A minority of courts have held that the plaintiff must show *both* consumer association *and* that this consumer association "impairs the distinctiveness" of the famous mark.<sup>11</sup> Here, Apple of Cupertino would need to show that the bar calls to mind the global brand and that this calling to mind somehow impairs the global brand's distinctiveness. We refer to this as the *association plus impairment* standard.

The split among courts is surprising, and not just because the statutory language appears clearly to call for the *association plus impairment* approach. In the 2003 case *Moseley v. V Secret Catalogue, Inc.*, the Supreme Court explicitly stated, albeit in dicta, that "[b]lurring" is not a necessary consequence of mental association."<sup>12</sup> The Court could not have been clearer: the mere fact that consumers associate the defendant's mark with the plaintiff's cannot, without more, establish blurring. And yet a majority of courts continue to accept evidence merely of association as sufficient evidence of blurring.

In this Article, we make two contributions to the current debate over the nature of dilution by blurring and what, empirically, must be shown to establish it in court. First, we report findings from a new set of experiments that confirm that the majority *mere association* approach is fundamentally deficient. In support of the *Moseley* Court's dictum (and common sense), these experiments demonstrate that even when consumers associate a junior mark with a famous senior mark, this association does not necessarily result in any impairment of the ability of the senior mark to identify its source and associations. The practical significance of our experimental results is that survey and other evidence that shows merely that consumers associate the defendant's mark with the plaintiff's

---

<sup>9</sup> See, e.g., *Wrenn v. Boy Scouts of America*, No. 03 Civ. 04057, 2008 WL 4792683, at \*7 (N.D. Cal. 2008); *Hershey Co. v. Art Van Furniture Inc.*, No. 08 Civ. 14463, 2008 WL 4724756, at \*14-15 (E.D. Mich. 2008); *New York Yankees Partnership v. IET Products and Services, Inc.*, Opposition No. 91189692, 2015 WL 2455162, at \*8 (T.T.A.B. 2015).

<sup>10</sup> See, e.g., *Visa Intern. Service Ass'n v. JSL Corp.*, 610 F.3d 1088, 1090 (9th Cir. 2010) (granting summary judgment for likely dilution through domain name through mere association standard); see also MCCARTHY, *supra* note 5, at § 24:120 (criticizing the *Visa* reasoning).

<sup>11</sup> See, e.g., *Huginin v. Land O' Lakes, Inc.*, 815 F.3d 1064, 1067 (7th Cir. 2016) (dismissing dilution claim for absence of harm showing); *Gap, Inc. v. G.A.P. Adventures Inc.*, 2011 WL 2946384, \*17 (S.D.N.Y. 2011) (likelihood of association absent proof of impairment did not make out dilution claim).

<sup>12</sup> 537 U.S. 418, 434 (2003) (addressing a prior antidilution statute).

senior mark cannot be judged sufficient to establish blurring. All courts should require, as a minority of courts currently do, that plaintiffs also show that the association impairs the distinctiveness of the plaintiff's mark.

Our second contribution addresses a new form of survey evidence taken from the marketing literature: experimental response time surveys. These surveys purportedly show both association and impairment and should therefore satisfy the *association plus impairment* standard. Though to our knowledge such surveys have not yet been used at trial, experimental response time studies represent the current state of the art in dilution research.<sup>13</sup> Such studies expose treatment subjects to advertisements that ostensibly dilute certain targeted marks (for example, an advertisement for a fictitious brand called HEINEKEN popcorn), while control subjects are exposed to no such advertisements. The studies then compare the speed and accuracy with which the treatment and control subjects link various marks, including the targeted marks, with their respective sources and associations. The studies report that treatment subjects take longer to match the targeted marks with their true sources and associations, and are also less accurate in doing so. Both effects are offered as evidence of impairment of distinctiveness—and more generally, *contra* Farley and other skeptics, as evidence that dilution by blurring is indeed real.<sup>14</sup> In fact, these studies remain the best, if not the only, scientific evidence that dilution by blurring ever occurs.

We show, however, that the experimental protocols used in these leading response time experiments are fatally flawed. In short, they used the wrong control. We report the results of a second set of experiments in which we expose *both* treatment subjects and control subjects to ostensibly diluting advertisements. The difference is that control subjects were exposed to ostensibly diluting advertisements for marks other than the targeted marks that were the focus of the test. When using this experimental design, we find that treatment and control subjects produce no significant differences in the accuracy or speed with which they linked various marks with their respective sources and associations. We suggest that when test subjects are exposed to *any* diluting stimuli in the form of a brand-product linkage that is unfamiliar to them (for example, NIKE toothpaste), test subjects become wary and cautious, which increases their response times and impacts their accuracy with respect to linking tasks. When this wariness or *surprise* is controlled for, no evidence of blurring is found.

Our findings here have both practical and theoretical significance. As a practical matter, if response time studies are used in litigation, they must use the correct control. They must expose treatment subjects to ostensibly diluting stimuli that target the mark at issue in the litigation and control subjects to such stimuli that target other marks that bear no relation to the mark at issue but which

---

<sup>13</sup> We are aware of at least one trademark infringement case in which the plaintiff presented the results of response time surveys in discovery as evidence of dilution, but the case ultimately settled before trial.

<sup>14</sup> See Farley, *supra* note 4.

may nevertheless be surprising or unfamiliar. As a theoretical matter, our findings demonstrate that the best (and only) empirical evidence we have that blurring even occurs consists of studies that are invalid. This is alarming. Congress passed the first federal antidilution law in 1995, which was a poorly-drafted failure, and replaced it entirely in 2006.<sup>15</sup> Many states have had state antidilution statutes for decades. Countless court opinions have addressed a dilution cause of action. Yet at present there is no persuasive empirical demonstration that dilution exists in fact. Congress and the courts have devoted an enormous amount of effort to regulating a harm that remains at best hypothetical.

Part I reviews surveys currently used in dilution litigation that show mere association and reports the results of our experiments demonstrating that such surveys do not necessarily show impairment of distinctiveness and, as such, do not show blurring. Part II reviews current response time studies and reports the results of our experiments showing that they are incorrectly designed and invalid. Part III briefly discusses the general implications of our findings for the question of what dilution is and whether it even occurs. Part IV concludes by considering future directions for research.

## **I. Mere Association and Blurring**

As we mentioned briefly above, a majority of courts that have considered an antidilution claim have adopted the *mere association* standard for establishing blurring. That is, they have accepted evidence that consumers merely associate the plaintiff's and defendant's marks as sufficient evidence to show that the defendant's mark "impairs the distinctiveness" of the plaintiff's mark. And in certain leading cases, courts have relied on survey evidence showing association to find blurring. In section I.A, we review this survey evidence and explain why, even on its own terms, this evidence fails to establish blurring. We then present in section I.B the results of a set of experiments that confirm the insufficiency of evidence of *mere association*.

### **A. Survey Evidence of Mere Association**

Those courts that apply the *mere association* standard generally rely on two forms of survey evidence as proof of blurring. We review them in turn.

#### **1. The Nikepal Survey Approach**

The first and most prominent form of survey evidence that *mere association* courts have accepted consists of variations on the survey method used by the plaintiff in *Nike, Inc. v. Nikepal Int'l, Inc.*<sup>16</sup> In *Nikepal*, the defendant used the mark NIKEPAL as the name of its business distributing glass syringes and other

---

<sup>15</sup> Pub. L. 109-312, §120 Stat. 1730 (codified as 15 U.S.C. §1125(c) and replacing the Federal Trademark Dilution Act, Pub. L. 104-98); see Brian A. Jacobs, *Trademark Dilution on the Constitutional Edge*, 104 Colum. L. Rev. 168, 174 (2004) (discussing the interpretive uncertainties of the repealed Federal Trademark Dilution Act of 1995).

<sup>16</sup> No. 05 Civ. 1468, 2007 WL 2782030 (E.D. Cal. Sept. 18, 2007).



laboratory products.<sup>17</sup> Nike conducted a telephone survey of the defendant's current and prospective customers in which it asked them about "their perception of a website called nikepal.com."<sup>18</sup> Specifically, the survey asked: "What if anything, came to your mind when I first said the word Nikepal?" Unsurprisingly, 87% of respondents stated that they thought of the plaintiff or its products. The survey expert, and the *Nikepal* court, took this as evidence of blurring.<sup>19</sup> Other courts have accepted the results of similar surveys as evidence of blurring.<sup>20</sup>

The *Nikepal* court did not explain how *mere association* translates into evidence that use of NIKEPAL impairs the distinctiveness of the NIKE mark. Numerous trademark commentators have criticized the *Nikepal* survey method as failing to present persuasive evidence of dilution,<sup>21</sup> and we believe these criticisms are valid. Nike is one of the world's best-known brand names. The fact that a consumer thinks of a famous mark when he or she sees a word containing that mark may not mean that the distinctiveness of the famous mark is "blurred" or harmed in any way. Indeed, because the association calls the famous mark to mind, its strength and salience may conceivably be reinforced.<sup>22</sup> Each time a Brooklynite goes into that Apple bar, he or she may be reminded that Apple, the iPhone maker, is a wonderful company. The measure used in the *Nikepal* case cannot tell us which of the outcomes is more likely, and for that reason lacks construct validity; i.e., it cannot be taken as a valid measure of harm.

In order to assess harm, we need a valid measure. The marketing literature is very clear on how a brand suffers harm. Marketers would describe harm to a

---

<sup>17</sup> *Id.* at \*2.

<sup>18</sup> *Id.* at \*4.

<sup>19</sup> *Id.*

<sup>20</sup> See, e.g., *Perfumebay.com, Inc. v. eBay, Inc.*, 506 F.3d 1165, 1172 (9th Cir. 2007) (discussing a similar telephone survey asking respondents what website or company they would think of if they encountered the name "BAY" on the Internet). See generally Krista F. Holt & Scot A. Duvall, *Chasing Moseley's Ghost: Dilution Surveys Under the Trademark Dilution Revision Act*, 98 TRADEMARK REP. 1311 (2008) (reviewing survey evidence of dilution considered by federal courts). But see *Starbucks Corp. v. Wolfe's Borough Coffee, Inc.*, 736 F.3d 198, 209–11 (2d. Cir. 2013) (3% response insufficient to prove actual association).

<sup>21</sup> See, e.g., Matthew D. Bunker & Kim Bissell, *Lost in the Semiotic Maze: Empirical Approaches to Proof of Blurring in Trademark Dilution Law*, 18 COMM. L. & POL'Y 375, 384 (2013) ("Aside from the problem of conflating association with dilution, the [Nike] survey certainly provides no evidence of dilutive harm since there is no baseline measurement of the strength of Nike's brand prior to Nikepal's entry into the marketplace.").

<sup>22</sup> Cf. *Louis Vuitton Malletier S.A. v. Haute Diggity Dog, LLC*, 507 F.3d 252, 267 (4th Cir. 2007) ("[B]y making the famous mark an object of the parody, a successful parody might actually enhance the famous mark's distinctiveness by making it an icon. The brunt of the joke becomes yet more famous.").

brand (or trademark) as a diminution of (customer-based) brand equity.<sup>23</sup> Keller characterizes brand awareness and brand image as the components of customer-based brand equity.<sup>24</sup> Brand image in turn is built from strong, favorable, and unique brand associations. Damage caused by blurring conceptually fits into this framework as a weakening of brand associations, a compromise of the uniqueness of the associations, or a reduction in the favorability of the associations the trademark owner has worked hard to cultivate.<sup>25</sup> The *Nikepal* approach cannot indicate that an association is tied to brand weakening, and as such is not probative of blurring.

## 2. Association and Confusion Survey Results

Courts have also relied on likelihood of confusion surveys as evidence of dilution by blurring. For example, in *Jada Toys, Inc. v. Mattel, Inc.*,<sup>26</sup> Mattel, which manufactured toy automobiles under the HOT WHEELS mark, argued that Jada Toys' sale of toy automobiles under the HOT RIGZ mark would both confuse consumers and blur Mattel's mark. Mattel proffered two confusion surveys.<sup>27</sup> The first exposed respondents to the HOT RIGZ mark and then asked a series of questions designed to determine who the respondents believed "puts out or makes" toy vehicles with that name.<sup>28</sup> The second survey exposed respondents to a HOT RIGZ package and asked similar questions.<sup>29</sup> In reversing the district court's partial grant of summary judgment to Jada Toys, the Ninth Circuit Court of Appeals pointed to both surveys as "significant evidence of actual association between the alleged diluting mark and the famous mark,"<sup>30</sup> and found that a "reasonable trier of fact could conclude that this evidence was

---

<sup>23</sup> See, e.g., Kevin Lane Keller, *Conceptualizing, Measuring, and Managing Consumer-Based Brand Equity*, 57 J. MARKETING 1 (1993).

<sup>24</sup> See Kevin Lane Keller, *Strategic Brand Management: Building, Measuring, and Managing Brand Equity* 48-51 (2013).

<sup>25</sup> This view follows the Associative Network Theory of memory in psychology. See Allen M. Collins & Elizabeth R. Loftus, *A Spreading-Activation Theory of Semantic Processing*, 82 PSYCH. REV. 407 (1975). That theory models information in (long-term) memory as networks of nodes connected by links. The nodes consist of trademarks and concepts linked to or associated with those marks. These concepts include product categories and attributes possessed. When exposed to a trademark, the network containing that trademark is activated and the activation signal spreads outward to the concepts linked to that node. Blurring presumes that that signal is no longer sufficiently strong to activate the node of the desired association.

<sup>26</sup> 518 F.3d 628 (9th Cir. 2008).

<sup>27</sup> *Id.* at 636. See also Bunker & Bissell, *supra* note 20, at 1332-35 (discussing the *Jada Toys* surveys in detail).

<sup>28</sup> *Jada Toys*, 518 F.3d at 636 (describing Mattel's confusion survey modeled after the *Eveready* design); see MCCARTHY, *supra* note 5, at § 32:174 (describing the *Eveready* design).

<sup>29</sup> *Jada Toys*, 518 F.3d at 636.

<sup>30</sup> *Id.*

sufficient to establish the existence of a likelihood of dilution.”<sup>31</sup> The Second Circuit Court of Appeals has more recently written approvingly of the *Jada Toys* court’s reasoning and reliance on confusion survey evidence as potential evidence of dilution, stating that “[s]ource confusion may be probative of association.”<sup>32</sup>

As with the *Nikepal* survey format, consumer confusion surveys show consumer association but they do not show blurring. Indeed, confusion surveys arguably show the *opposite* of blurring. Recall that blurring describes situations in which, due to the similarity of the plaintiff’s and the defendant’s marks, consumers see the plaintiff’s mark and must think for a moment to determine whether the mark is referring to the plaintiff or the defendant—APPLE for the high-technology company or for the bar in Brooklyn. In this scenario, consumers are not confused as to source. They know that APPLE may refer to one of two different companies. By contrast, consumer confusion surveys expose situations in which consumers believe that both marks, due to their close similarity, refer only to the plaintiff. In this sense, the defendant’s mark reminds consumers of and reinforces the link between the plaintiff’s mark and the plaintiff. Consumers do not think that there are two different companies. They think that there is only APPLE, the famous high-technology company that has now apparently expanded its brand into cocktail bars.

In other words, in the case of consumer confusion, on encountering the defendant’s mark, the confused consumer associates it with the plaintiff’s mark and believes that it originates in the plaintiff. By contrast, in the case of blurring, on encountering the defendant’s mark, the blurred consumer associates it with the plaintiff’s mark but knows that it originates in the *defendant*. Evidence of consumer association that leads to consumer confusion is evidence that consumers think there is only one company rather than two; those confused consumers cannot experience blurring because they associate both marks with the *same* company.

On these grounds, many commentators have been critical of courts’ conflation of evidence of association that leads to confusion with evidence of association that leads to blurring.<sup>33</sup> This conflation risks merging confusion and dilution, which are two very different causes of action. Yet plaintiffs, it seems, are only too eager to confuse the matter and, having conducted a confusion survey, present its results as evidence of blurring as well.

---

<sup>31</sup> *Id.*

<sup>32</sup> *Starbucks Corp. v. Wolfe’s Borough Coffee, Inc.*, 736 F.3d 198, 211 n.14 (2d Cir. 2013).

<sup>33</sup> *See, e.g., MCCARTHY, supra* note 5, at § 24:116 (The statutory [dilution] requirement of ‘association’ is not the same kind of mental link that must occur for classic trademark infringement by a likelihood of confusion.”)

## **B. Survey Evidence of Mere Association is Not Evidence of Impairment**

We believe, as the *Moseley* Court stated, that mere association evidence fails on its own terms to establish blurring. But we also present experimental evidence confirming that even when consumers associate a junior mark with a senior famous mark, this does not result in the weakening of the association of the famous mark with its source or other associations. Furthermore, such consumer association does not necessarily result in any material change in consumers' purchasing preferences. *Mere association* on its own does not damage the senior mark.

### **1. Brand Selection Pretest**

Because we intended in these experiments to study blurring in the form of changes in brand associations, we sought out brands with strong, relatively well-defined associations. We chose automobiles as a product category and collected data from 500 people to identify brands and associations that could potentially be blurred.<sup>34</sup> The research began by showing respondents the names of up to twenty popular car brands. The brand names were shown one by one in random order. For each brand name, respondents were asked whether they were familiar with the brand. Familiarity with the brand was defined as knowing something about it other than that it exists. Respondents responded “yes” or “no.”

As soon as a respondent identified five familiar brands, we stopped showing that respondent brands and moved on to the next phase. In this second phase, respondents were asked to name the top five things that came to mind when they think of each of the brands they identified as familiar; i.e., we asked for each individual's top five brand associations. Our assumption was that very strong brands would bring to mind a relatively concentrated set of associations, most of which would be positive. Conversely, weaker brands would bring to mind a larger set of more diffuse, and sometimes negative, associations.

To determine how concentrated a brand's associations were, we collected all of the associations that respondents provided for each brand. We then grouped like associations and calculated the percentage of subjects identifying each group. We then added up the percentages for the five most common association groups mentioned for each brand. This gave us a concentration score. The mean concentration score was 39 percent. We considered a brand “strong” if its concentration score was at least one standard deviation above the mean.

This left us with Mercedes (57.41 percent), BMW (55.92 percent), and Infiniti (48.34 percent) as our three strongest brands. Coincidentally, all three brands are luxury car brands. Respondents tended to emphasize this fact most when asked to list words they associate with each mark. In all three cases, the most common associations named were characteristics like “luxury,” “expensive,” as well as words that connote wealth and high socioeconomic

---

<sup>34</sup> We recruited subjects for our studies on Amazon's Mechanical Turk platform. We discuss Mechanical Turk in Appendix A.

status. Respondents also associated words like “fast” and “good looking” with the Mercedes, BMW, and Infiniti brands, but to a lesser extent.

We first selected MERCEDES because it had the highest concentration score. For our second test brand, we chose INFINITI. INFINITI had a relatively high concentration of positive associations, but it was among the least familiar brands in our sample. In fact, only 41.1 percent of respondents who saw an INFINITI prompt said they knew something about the brand other than that it existed. By comparison, 60.8 percent of respondents who saw a Mercedes prompt and 63.8 percent of respondents who saw a BMW prompt said that they knew something about those brands. As such, although INFINITI has relatively clear associations among those who are familiar with it, it may be more easily diluted since its associations are not as widely held.

## **2. Study 1: Examining Brand Association Strength**

Study 1 tested whether blurring advertisements affected the strength of the association between the MERCEDES and INFINITI brands and both their product category (cars) and the top associations (wealth, luxury) previously found for each brand in our initial brand selection pretest. A total of 2012 subjects participated in the experiment. Each was randomly assigned to a brand (MERCEDES or INFINITI), then randomly assigned to a treatment or control group. The 503 subjects in the INFINITI control group and the 506 subjects in the MERCEDES control group saw three “filler” text-only “tombstone” advertisements for well-known brands (e.g. PERRIER sparkling water, CHASE bank, UNITED VAN LINES) that contained true representations of each brand’s product category and attributes. Subjects in our two treatment groups saw the three filler ads plus an ad for a fictitious diluting brand. The 505 subjects assigned to the MERCEDES treatment group saw an ad for a diluting brand called “MERCEDES Toothpaste” and the 498 subjects assigned to the INFINITI treatment group saw an ad for a diluting brand called “INFINITI Toothpaste.”<sup>35</sup>

All subjects were then shown twenty-three brand-word pairs and told to state the degree to which they associate the brand name and the word on a five-point Likert scale,<sup>36</sup> where a score of 1 means they associate the brand and the word “a great deal” and a score of 5 means they associate the brand and the word “not at all.” Eighteen of these pairs were filler pairs for brands unrelated to cars, such as CHASE-money, CHASE-food, COCA COLA-thirst, COCA COLA-expensive, NIKE-slow, NIKE-poor, UNITED-plane, and UNITED-boxes.

---

<sup>35</sup> The tombstone advertisements used in the experiment are reproduced in Appendix B.

<sup>36</sup> Likert scaling is a unidimensional multi-point scaling method that is commonly used to measure agreement with a proposition or (as in our experiment) the degree of association between two identified things or propositions. Unlike binary “yes or no” measures of agreement, Likert scales, which commonly measure agreement or association along a 5-point or 7-point scale, provide more information about the strength of the agreement or association that is being measured. *See* Donald R. Lehmann et al., *Marketing Research* 143 (1998).

Randomly mixed in with the eighteen filler pairs were five brand-word pairs for one of our target marks. The pairs for our target marks included the product category “cars” and two words that capture the high-status qualities that respondents in our brand selection survey tended to associate with both marks: “luxury” and “wealth.” Finally, we included two words that are not associated with the MERCEDES and INFINITI car brands: the product category “toothpaste” and the attribute “cheap.” Subjects in the MERCEDES group therefore saw the pairs MERCEDES-cars, MERCEDES-wealth, MERCEDES-luxury, MERCEDES-toothpaste, and MERCEDES-cheap randomly distributed among the eighteen filler pairs. Subjects in the INFINITI group saw INFINITI-cars, INFINITI-wealth, INFINITI-luxury, INFINITI-toothpaste, and INFINITI-cheap mixed in with the eighteen filler pairs.

If the MERCEDES and INFINITI Toothpaste ads were causing dilution by blurring, we would expect that subjects in the treatment groups exposed to such ads would show two differences from subjects in the control group. First, as compared to control subjects, treatment subjects would more strongly associate the car brands with the product category “toothpaste.” Second, treatment subjects would less strongly associate the car brands with the product category “cars.” They may also become less likely to associate these marks with words that our pretests indicated to be strongly associated with MERCEDES and INFINITI, such as “luxury” and “wealth,” and more likely to associate to these brands with the word “cheap,” which better describes an ordinary, low-cost product like toothpaste than an expensive luxury automobile. Such differences would indicate that distinctiveness has been impaired.

The results presented in Table 1 support only part of this story, however. First, our results show that our blurring stimulus produced a new association – i.e., an association between MERCEDES or INFINITI, and toothpaste – for a significant number of subjects. Specifically, the average effect of the diluting toothpaste ad versus control was 0.369 of a point on a 5-point scale in the case of MERCEDES and 0.73 of a point on a 5-point scale in the case of INFINITI. Both effects are statistically significant ( $p < 0.01$ ), and they remained significant in regressions that controlled for subjects’ stated level of familiarity with the MERCEDES and INFINITI brands, which we asked about at the end of the interview.

Crucially, though, Table 1 also shows no evidence that this new association between the test brands and the toothpaste product category is accompanied by a *weakening of the association* between the test brands and their true product category or their principal product attributes. Subjects who saw the MERCEDES Toothpaste and INFINITI Toothpaste ads were no less likely than control group subjects to associate target marks with the product category “cars” or attributes like “luxury” and “wealth.” Subjects who saw the MERCEDES Toothpaste and INFINITI Toothpaste ads were furthermore no more likely than control group subjects to associate these brands with the word

**Table 1. Likert Association Strength Tests. Difference between Treatment Group Mean and Control Group Mean.**

<b>Treatment Group: Mercedes/Infiniti Toothpaste Ad</b>				
<b>Control Group: No 4th Ad</b>				
<i>Mercedes-Luxury</i>	<i>Mercedes-Wealth</i>	<i>Mercedes-Cars</i>	<i>Mercedes-Toothpaste</i>	<i>Mercedes-Cheap</i>
0.052 (0.050)	0.019 (0.059)	0.040 (0.040)	-0.369*** (0.048)	-0.016 (0.032)
<i>Infiniti-Luxury</i>	<i>Infiniti-Wealth</i>	<i>Infiniti-Cars</i>	<i>Infiniti-Toothpaste</i>	<i>Infiniti-Cheap</i>
-0.009 (0.078)	0.124 (0.081)	0.055 (0.075)	-0.730*** (0.058)	-0.029 (0.043)

Standard Errors in parentheses.

Two tailed test: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

“cheap.” In short, Table 1 shows no evidence of distinctiveness having been impaired.

In sum, the MERCEDES and INFINITI Toothpaste ads contributed to the formation of a new association between the targeted brands and a new product category. Yet at the same time that we observed the formation of a new association among subjects in the treatment groups, we found no evidence in this experiment that the creation of this new association was accompanied by the blurring or weakening of pre-existing associations, either with respect to product category or product attributes. Comparing average responses, there were no statistically significant differences between the degree to which subjects in the treatment and control groups associated certain words with the target brands.

### **3. Study 2: Dilution at the Cash Register**

Study 1 shows that new associations do not necessarily lead to the weakening of other associations, but the ultimate question as to whether distinctiveness is impaired is whether these new associations have some effect on the “selling power” of the famous brand. Study 2 engaged this issue. It tested whether ostensibly blurring advertisements and the new associations they produce affect consumer preferences and consumers’ intent to purchase the targeted brand.

**Table 2. Preference Ranking Experiments. Difference between Treatment Group Mean and Control Group Mean.**

Treatment Group: Mercedes/Infiniti Toothpaste Ad	
Control Group: Nike Ad	
Mercedes	Infiniti
0.182	-0.038
(0.169)	(0.242)

Standard errors in parentheses.

One tailed test: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Study 2 consisted of a preference-ordering protocol in which 1008 subjects were first shown up to twenty car brands in random order and asked to state whether or not the brand was familiar. When five familiar brands were selected, respondents moved to the next stage of the experiment. Half of the subjects who said they were familiar with either MERCEDES or INFINITI were shown a diluting MERCEDES Toothpaste or INFINITI Toothpaste ad in addition to three filler ads (treatment group). The other half were shown a NIKE Toothpaste ad in addition to the three filler ads (control group).<sup>37</sup>

After viewing the ads, subjects were asked to rank the five brands they had identified as familiar in order of preference, from 1 (favorite brand) to 5 (least favorite). Overall, the mean rank for MERCEDES was 2.31 and mean rank for INFINITI was 2.89.

Table 2 shows the differences between the mean preference ranking for treatment and control group subjects for each brand. In both cases, the differences are not statistically different from zero. We are unable to find any evidence from this protocol that the diluting ad caused subjects to rank MERCEDES or INFINITI lower.

\* \* \*

To recall, federal trademark law states that “‘dilution by blurring’ is association arising from the similarity between a mark or trade name and a famous mark that *impairs the distinctiveness* of the famous mark.”<sup>38</sup> This statutory language implies that some associations do not “impair the

<sup>37</sup> For reasons we explain more fully in the next Part, the control group was shown a toothpaste ad for a famous brand other than MERCEDES or INFINITI to ensure that control subjects were exposed to some ostensibly diluting advertisement, though one not directed at the targeted brands.

<sup>38</sup> 15 U.S.C. § 1125(c)(2)(B) (emphasis added).



distinctiveness” of the plaintiff’s mark. Otherwise, if all association did result in impairment, then there would be no need to add the limiting phrase “that impairs the distinctiveness of the famous mark.” The statutory scheme assumes an underlying fact: association does not necessarily lead to impairment.

Studies 1 and 2 confirm this assumption—and the *Moseley* Court’s dictum. We exposed subjects to diluting stimuli that created new associations between the targeted brands and new products and brand characteristics in these subjects’ shared associative memory networks.<sup>39</sup> But despite the new associations, these subjects showed no weakening of the association between the targeted brands and their traditional products and characteristics. Nor did these subjects reveal any change in purchasing preferences. On this basis, we conclude that courts that have treated *mere association* as tantamount to proof of the blurring form of trademark dilution have erred. Plaintiffs must present something more than evidence of association to establish the likelihood that a defendant’s mark will suffer dilution via blurring.

## II. Response Time Experiments and Blurring

Association surveys fail to establish blurring, but they are not the only tool available to litigants and courts. Over the last two decades, social science research has developed more sophisticated tests for blurring based on *response time measurements*, and these tests have begun to penetrate ongoing litigations.<sup>40</sup> Unlike association surveys, response time tests purportedly show disassociation in that subjects exposed to ostensibly diluting stimuli appear to take longer than control subjects to link targeted marks with their traditional product categories and product attributes and characteristics. In this Part, however, we show that previous versions of these tests were flawed. When response times tests use the proper control, no evidence of dilution is found. We first review these previous experiments. We then report the results of our own response time experiments and explain their significance.

### A. Previous Response Time Experiments

We focus on two response time studies: Maureen Morrin and Jacob Jacoby’s pathbreaking 2000 study *Trademark Dilution: Empirical Measures for an Elusive Concept*,<sup>41</sup> and Chris Pullig, Carolyn J. Simmons, and Richard G. Netemeyer’s influential 2006 study *Brand Dilution: When Do New Brands Hurt Existing Brands?*<sup>42</sup>

---

<sup>39</sup> See Collins & Loftus, *supra* note 24.

<sup>40</sup> See *supra* note 13.

<sup>41</sup> Maureen Morrin & Jacob Jacoby, *Trademark Dilution: Empirical Measures for an Elusive Concept*, 19 J. PUB. POLY’ & MKTG. 265, 268 (2000) [hereinafter Morrin & Jacoby].

<sup>42</sup> Chris Pullig et al., *Brand Dilution: When Do New Brands Hurt Existing Brands?* 70 J. Mktg. 52 (2006) [hereinafter Pullig, Simmons, & Netemeyer]. For other dilution experiments, see Maureen Morrin et al., *Determinants of Trademark Dilution*, 33 J. CONSUMER RES. 248 (2006); Hanelie Kruger & Christo Boshoff, *The Influence of*

## 1. Morrin and Jacoby

Morrin and Jacoby set out to test the hypothesis that “[c]onsumers exposed to trademark-diluting advertisements will (a) commit more brand recognition errors and (b) exhibit slower brand recognition reaction times than will consumers not exposed to such advertisements.”<sup>43</sup> Their method was straightforward (and served as a model for our own Pilot Studies 1 and 2 and Study 3, all discussed below). They exposed sixty-four subjects each to a total of six “tombstone” text advertisements, some of which were ostensibly diluting of three target brands (HEINEKEN, HYATT, and GODIVA), and then recorded the speed and accuracy with which the subjects attempted to match the target brands with their product categories and characteristics.<sup>44</sup>

The treatment group consisted of thirty-three subjects who viewed three text-only “filler” advertisements for well-known brands that contained true representations of each brand’s product category and attributes (for CHASE MANHATTAN, KODAK, and PERRIER), one blurring advertisement, one tarnishing advertisement,<sup>45</sup> and one brand extension advertisement.<sup>46</sup> The blurring advertisement took the form of an advertisement either for HEINEKEN popcorn or HYATT legal services. Each blurring advertisement contained a disclaimer stating either that “HEINEKEN popcorn is NOT associated with the makers of Heineken beer” or that “HYATT legal services is NOT associated with Hyatt Hotels, Inc.”<sup>47</sup> The tarnishing advertisement was for DOGIVA dog biscuits and contained no disclaimer with respect to GODIVA. The brand extension advertisement was for either HEINEKEN popcorn or HYATT legal services—whichever was not used in the blurring advertisement—and apparently contained language announcing the brand extension.

One control group consisted of sixteen subjects who were exposed to the three filler advertisements and to three unrelated advertisements. A second control group consisted of fifteen subjects who were exposed to the three filler

---

*Trademark Dilution on Brand Attitude: An Empirical Investigation*, 24 MANAGEMENT DYNAMICS 50 (2015).

<sup>43</sup> Morrin & Jacoby, *supra* note 39, at 268.

<sup>44</sup> *Id.* at 268-70.

<sup>45</sup> Though Morrin & Jacoby focused their study on dilution by blurring, they also tested for dilution by tarnishment. A senior trademark is tarnished when a junior mark portrays the senior mark in a negative light – for example, by associating it with illegal activities or sexually suggestive messages. The theory underlying liability for dilution by tarnishment is that, even in the absence of confusion regarding the source of the senior mark owner’s products, the selling power of the senior mark may be harmed by the unsavory association. *See MCCARTHY*, *supra* note 5, at § 24:70.

<sup>46</sup> That is, an advertisement suggesting that the targeted brand was extending its reach into a product category different from the one with which it had been traditionally associated by consumers.

<sup>47</sup> Morrin & Jacoby, *supra* note 39, at 275.

advertisements and three reinforcing advertisements for HEINEKEN beer, HYATT hotels, and GODIVA chocolate.<sup>48</sup>

Immediately after viewing the six advertisements, subjects engaged in a computer task in which two words appeared sequentially on a screen. Subjects were instructed to indicate as quickly but as accurately as possible whether the two words “represented a match,”<sup>49</sup> by hitting one key for “Yes” or another for “No.” Among the word pairs presented were pairs relating to category membership (for example, GODIVA/chocolates, chocolates/GODIVA) and attribute possession (for example, GODIVA/rich taste, rich taste/GODIVA).

Morrin and Jacoby reported that the mean response times of the treatment group were significantly higher than those of either of the control groups, but only with respect to matches relating to HEINEKEN and GODIVA and not with respect to matches relating to HYATT. For matches relating to HEINEKEN and GODIVA taken together, treatment subjects’ mean response time was 836 milliseconds while it was 672 milliseconds for the control group exposed to reinforcing advertisements and 713 milliseconds for the control group exposed to unrelated advertisements.<sup>50</sup> Meanwhile, for matches relating to HYATT, the treatment group’s mean response time was 680 milliseconds, which was exactly the same as the reinforced control group’s mean response time and substantially less than the unreinforced control group’s mean response time of 810 milliseconds.<sup>51</sup> Morrin & Jacoby further reported that the treatment group’s matching accuracy was lower than both control groups and that this difference was statistically significant. Specifically, the treatment group accurately judged 73.1% of the matches, while the unreinforced control group and the reinforced control group accurately judged 84.3% and 88.9%, respectively, of the matches.<sup>52</sup>

As Morrin and Jacoby saw it, these results fully supported the hypothesis that consumers exposed to diluting advertisements will commit more brand recognition errors, but only partially supported the hypothesis that such consumers will react more slowly to brand recognition tasks. They speculated that subjects’ increased familiarity with the HYATT brand moderated the effect on response times of the HYATT legal services advertisement. The authors conducted a subsequent study that tested consumers’ ability after exposure to diluting stimuli to recall the categories of goods in connection with which certain

---

<sup>48</sup> *Id.* at 268.

<sup>49</sup> *Id.* at 269.

<sup>50</sup> *Id.*

<sup>51</sup> *Id.*

<sup>52</sup> *Id.* Morrin & Jacoby present some results on recognition accuracy in addition to reaction time. See Morrin & Jacoby, *supra* note 39, at 269. However, they do not detail the associations that they examine apart from brand category (e.g. GODIVA chocolate). As such, it is difficult to know what “accuracy” means. Finally, the numbers presented in their Table 1 (top panel) are inconsistent. *Id.* Unlike the bottom panel of the table, the mean row is not equal to the mean of the rows. As such, it is difficult to interpret or accept the findings.

brand were used. The study tested both well-known and unknown brands. Based on the results of this study, Morrin and Jacoby concluded, “it appears that very strong brands are immune to dilution because their memory connections are so strong that it is difficult for consumers to alter them or create new ones with the same brand.”<sup>53</sup> This conclusion, if correct, carries with it significant public policy implications since federal antidilution law is designed to protect only “famous” brands (i.e., those that are “widely recognized by the general consuming public of the United States.”)<sup>54</sup>

## 2. Pullig, Simmons, and Netemeyer

Now armed with a methodological paradigm, subsequent response time research addressed the logical next question: under what conditions would we expect dilution to be more likely to occur? Pullig, Simmons, and Netemeyer built on the Morrin & Jacoby approach to address this question by investigating potential moderators of trademark dilution.<sup>55</sup> In particular, they were interested in whether the similarity between the categories, attributes, or benefits of the senior and junior users moderated the threat of dilution by blurring.

Pullig, Simmons, and Netemeyer’s experimental protocol was complex and need not be described in detail here. In essence, they focused on BIG RED, a brand of chewing gum reasonably well-known in the U.S. for having a cinnamon flavor, and used a hypothetical BRAND A as a control. To study the effect of category similarity, they exposed eighty-nine subjects to a total of three advertisements according to a “2 (category similarity: high [bubble gum] or low [snack bars]) x 2 (new product: junior brand [BIG RED] and unbranded [BRAND A]) x 2 (order of ad presentation) between subjects design.”<sup>56</sup> More specifically, one advertisement was for bubble gum (branded as either BIG RED, BRAND A, or BAZOOKA), one was for bite-size candy (branded as SKITTLES), and one for snack bars (branded as either BIG RED, BRAND A, or NUTRI-GRAIN). Each advertisement contained the brand name, a product category, and two claims. To hold constant attribute similarity, certain bubble gum and snack bar ads claimed “great cinnamon flavor.”<sup>57</sup> The BIG RED bubble gum advertisement also promised to “blow awesome bubbles,” while for the BIG RED snack bars advertisement, the other claim stated “enriched with vitamins and minerals.”<sup>58</sup>

After viewing the advertisements, subjects engaged in a computer-administered task. The subject was presented with a brand name or attribute, and, after pressing the space bar, an attribute or brand name (depending on what was not presented first) appeared. The subject then pressed a key as quickly but as

---

<sup>53</sup> *Id.* at 274.

<sup>54</sup> 15 U.S.C. § 1125(c)(2)(A).

<sup>55</sup> *See* Pullig, Simmons, & Netemeyer, *supra* note 40.

<sup>56</sup> *Id.* at 55-57.

<sup>57</sup> *Id.* at 57.

<sup>58</sup> *Id.*

accurately as possible to indicate whether the brand name and attribute matched.<sup>59</sup>

Pullig, Simmons, and Netemeyer reported that subjects exposed to the use of BIG RED in the similar product category (bubble gum) exhibited significantly faster response times linking BIG RED and chewing gum and linking BIG RED and cinnamon flavor than subjects exposed to BRAND A bubble gum. The difference in speed ranged from 174 milliseconds to 302 milliseconds. These results suggested that junior users in a similar product category reinforce rather than blur the senior mark.<sup>60</sup>

By contrast, subjects exposed to the use of BIG RED in the different product category (snack bars) exhibited significantly slower response times than subjects exposed to BRAND A snack bars. Here the difference ranged from 119 milliseconds to 334 milliseconds. These results suggested to the authors that junior users in a different product category blur the mark.<sup>61</sup>

In a second study, Pullig, Simmons, and Netemeyer manipulated both category and attribute similarity. Their protocol was based on that of the previous study but with an additional dimension varying attribute similarity between high, “cinnamon flavor and breath freshening,” and low, “strawberry flavor and containing immunity-boosting Echinacea.” Their response time results suggested to them that a junior user in a similar product category with similar product attributes reinforces the senior user’s mark. On the other hand, a junior user in a different product category with different attributes blurs the senior mark. However, the response time results were otherwise mixed and suggest that attribute similarity may sometimes offset category dissimilarity and forestall blurring.<sup>62</sup>

## **B. Response Time Pilot Studies: Are We Really Observing Dilution?**

A leading criticism of the Morrin & Jacoby study in particular and to some extent of the Pullig, Simmons, and Netemeyer study is that they failed to place their test subjects in a purchasing context. Rebecca Tushnet has persuasively argued that consumers are sophisticated enough to disambiguate terms in light of context.<sup>63</sup> For example, at an airport, when someone refers to “United” or “Delta,” no one must “think for a moment” to know that the references are to airlines and not to the Manchester soccer club or the faucet maker. Even in a setting, such as online search, where there is less context, consumers are often able to disambiguate various uses of a mark using the information that

---

<sup>59</sup> *Id.*

<sup>60</sup> *Id.* at 58.

<sup>61</sup> *Id.*

<sup>62</sup> *Id.* at 60.

<sup>63</sup> Rebecca Tushnet, *Gone in Sixty Milliseconds: Trademark Law and Cognitive Science*, 86 TEX. L. REV. 507 (2008).

## Airline Tickets & Flights: Book Direct with Delta Air Lines - Official Site

<https://www.delta.com/> ▼

Delta Air Lines, a leader in domestic and international travel, offers airline tickets & flights to over 300 destinations in 60 countries. Book direct at [Delta.com](https://www.delta.com/).

### Flight Status

flight status description.

### My Delta

SkyMiles - Reset Password - Forgot Your Login? - ...

### Advanced Search

Delta Air Lines flies to over 300 destinations in 60 countries ...

### Skip to My Trips

My Trips · Book a Trip · Flight Status · Check In · Sign Up. Log ...

### My Trips & Check-In

MY TRIPS & CHECK-IN. Know the ins and outs of checking in to ...

### Talk To Us

Delayed, Lost or Damaged Baggage. Making sure your ...

[More results from delta.com »](#)

## Delta Faucet | Bathroom & Kitchen Faucets, Showers, Toilets, Parts ...

<https://www.deltafaucet.com/> ▼

The **Delta** Faucet brand delivers well-made, stylish faucets, shower heads and other kitchen and bathroom accessories. Browse now to customize your kitchen or bathroom.

accompanies search results. Pictured above, for example, are results from a Google search for “Delta.” Using the search test, one can clearly differentiate the airline from the faucet maker.

Blurring, in other words, is possible in theory but rarely if ever likely in practice, because of the disambiguating effect of purchasing context. We initially set out to determine whether placing treatment and control subjects in a purchasing context would affect the comparative speed and accuracy with which they link marks with their sources and associations. What we found led us to recognize that the basic design of previous response time studies was flawed.

### 1. Pilot Study 1: The Surprise Effect

We began our response time study by conducting a small-sample pilot to pretest our survey instrument. The objective of this pilot was to measure the speed and accuracy with which subjects identified whether the terms MERCEDES-CARS and INFINITI-CARS are a match. One hundred subjects were randomly assigned to a mark—either MERCEDES or INFINITI—and then to one of three conditions within each mark. Each mark-condition combination had between fifteen and seventeen respondents. The pilot protocol initially followed the method that we used in Study 1. Subjects in the control condition saw the three filler text-only “tombstone” advertisements for well-known brands (PERRIER sparkling water, CHASE bank, UNITED VAN LINES), while

subjects in our two treatment groups saw the three filler ads plus either the ad for “MERCEDES Toothpaste” or the ad for “INFINITI Toothpaste.”

After looking at the ads, subjects in one of the treatment groups participated in a brief car-buying role-play task—i.e., a task related to the targeted brands and intended to contextualize the matching task. These subjects were asked to write about the steps they would take in the process of purchasing a car, to think about how many dealerships they would visit, and to rank various features such as speed, roominess, and affordability. Subjects in the control group and the other treatment group performed a similar task after seeing the ads, but instead of being asked to imagine the process of buying a car, they were asked to imagine the process of searching for a job—a process unrelated to the targeted marks. In addition to allowing us the opportunity to expose some subjects to a purchasing context, these exercises served as distractors between the ads and the response time portion of the study. And if relevant context assists consumers in disambiguating the marks, we should see less, or no, response time delays for our subjects who were exposed to the car-buying context (relevant to the targeted marks), versus subjects exposed to the job-search context (irrelevant to the targeted marks).

In the response time portion of the protocol, which came next, all subjects were informed that on the following screens they would see a brand name and a word. They were told to click a “Yes” button if the mark and the word matched and click a “No” button if they did not. Subjects were asked to click as quickly as possible without sacrificing accuracy.

Subjects were then shown thirty brand-mark pairs. Twenty-five of these were “filler” pairs for marks such as MCDONALD’S, CHASE, COCA COLA, PERRIER, MICROSOFT, and UNITED. Mixed in with the filler pairs were pairs for one of our target marks, either MERCEDES or INFINITI. These pairs included the product category “cars” and two of the strongly-associated product attributes found for each mark in the Brand Selection Pretest (i.e., “wealth” and “luxury” in the case of Mercedes; “fast” and “luxury” in the case of Infiniti). As in our association studies (Studies 1 and 2, described above), we also included a product attribute that is not strongly associated with either mark (“cheap”) and a product category that is unrelated to cars (“printers”). Subjects in the MERCEDES group therefore saw MERCEDES-cars, MERCEDES-wealth, MERCEDES-luxury, MERCEDES-printers, and MERCEDES-cheap mixed in with the 25 filler pairs. Subjects in the INFINITI group saw INFINITI-cars, INFINITI-fast, INFINITI-luxury, INFINITI-printers, and INFINITI-cheap mixed in with the 25 filler pairs. Through Qualtrics, we were able to measure the time it took each subject to click “Yes” or “No,” down to one one-thousandth of a second.

When we compared average response times, our results for the MERCEDES group matched what others have found before: subjects who saw a diluting MERCEDES Toothpaste ad were slightly slower (by about two-tenths of a second) than control subjects to respond when the pair MERCEDES-cars flashed onto the screen. That delay was statistically significant (difference =

0.194 seconds;  $p < 0.05$ ). Subjects who saw the MERCEDES Toothpaste ad were also slower to respond when the pairs MERCEDES-printers (difference = 0.183;  $p < 0.05$ ) and MERCEDES-wealth (difference = 0.345;  $p < 0.05$ ) appeared on screen. We saw no such differences for the INFINITI pairs.

But we also noticed something that we did not expect. In addition to a statistically significant difference in average response time between treatment and control on the MERCEDES-cars, MERCEDES-printers, and MERCEDES-wealth pairs, *we noticed statistically significant differences between control and treatment groups on many of the filler pairs*. Subjects in the treatment groups, for example, were slightly slower to click “Yes” or “No” when pairs like MCDONALD’S-hamburgers (difference = 0.378;  $p < 0.01$ ), MCDONALD’S-affordable (difference = 0.203;  $p < 0.05$ ), MCDONALD’S-healthy (difference = 0.251;  $p < 0.01$ ), MICROSOFT-software (difference = 0.237;  $p < 0.05$ ), COCA COLA-expensive (difference = 0.211;  $p < 0.05$ ), and UNITED-boxes (difference = 0.416;  $p < 0.05$ ) appeared on screen.

The fact that our MERCEDES Toothpaste ad led to slower response times not only on the MERCEDES-cars pair, but in a total of 16 mark-word pairs completely unrelated to the MERCEDES mark, gave us pause and caused us to question what these response time experiments are actually measuring. Were we measuring dilution of the MERCEDES and INFINITI marks? Or was the diluting ad so unexpected by subjects that it distracted or surprised them and either caused them to suspect that experimenters were trying to trick them or simply diverted their attention from the task at hand? The latter would merely be an experimental artifact that could not be construed as harming the mark. We designed a second pilot to examine these issues.

## **2. Pilot Study 2: Controlling for Potential Surprise**

Our first pilot led us to hypothesize that an unexpected or surprising stimulus could give subjects pause as they proceed through the experimental task, resulting in slower response times in general, even for marks included in the task that the stimulus did not target. If this were the case, differences in response times between treatment and control groups would not be a sign of dilution, but rather an artifact of an experimental design that fails to properly control for the potentially distracting effects of any unexpected stimulus. Studies reported in the psychology and consumer studies literature support this understanding.<sup>64</sup>

---

<sup>64</sup> While no one had contemplated a surprise effect for such incongruent ads in the trademark literature, such an effect is consistent with prior work on response times and matching tasks. The psychology literature is replete with studies of perceptual and cognitive matching tasks. *See, e.g.,* Robert W. Proctor, *A Unified Theory for Matching-Task Phenomena*, 88 PSYCH. REV. 291 (1981). Such tasks generally involve either simultaneous or sequential presentation of stimuli (e.g. letters, numbers, words) that subjects judge to be the same or different according to criteria set by the experimenter.

One branch of this extensive literature involves priming. In these studies a priming stimulus, to which the subject generally does not have to react, is presented prior to another stimulus to which the subject generally does have to react. A neutral prime condition serves as a baseline against which a particular type of prime produces



Accordingly, we designed a second pilot to isolate this potential source of methodological error.

Pilot 2 was virtually identical in structure to Pilot 1, except that the protocol attempted to control for any unexpected stimulus. Again, we administered the study to 100 subjects. The treatment groups were the same as in Pilot 1. The sample sizes for the mark-treatment combinations ranged between fifteen and eighteen per group. Now, however, our control group subjects would see an ad for a fictitious NIKE Toothpaste in addition to the three filler ads. Our logic was that if our MERCEDES/INFINITI Toothpaste ads were causing subjects to slow down on the matching exercise simply because they were unexpected, that effect could be offset by including an ad that was just as unexpected in the control group; i.e. we would add a control ad to the control group. Statistically significant differences between treatment groups and the control group, in other words, should diminish or vanish if a surprise-type effect is at work.

When we compared average response times for the treatment and control groups we found that this is exactly what happened. Virtually no statistically significant differences in response times for the MERCEDES pairs, the INFINITI pairs, or the filler pairs appeared in Pilot 2. Across all groups and marks, there were twenty-nine statistically significant delays in Pilot 1 and only four statistically significant delays in Pilot 2. There were six significant delays on the Mercedes matches in Pilot 1 and none in Pilot 2 and none for Infiniti in either study. Also, comparison across treatments revealed no effect of context.

The results of our first two pilots led us to question the validity of using the traditional response time experiments to measure dilution. Whereas before we assumed that response time experiments were a valid way to measure dilution, we now believed that statistically significant results in the prior Morrin and Jacoby and Pullig, Simmon, and Neteymer experiments could have been due to failure to control for the “surprise” effect caused by the supposedly diluting ad. It could be that the mere presence of *any* unexpected mark-category combination

---

facilitation or inhibition in processing a test stimulus. The pilot studies reported in this paper as well as in the prior trademark dilution literature reviewed earlier all fall into this category.

Delays in response time are often evident when the prime is unrelated to the subsequent stimulus, a concept termed prime validity. James H. Neeley, *Semantic Priming and Retrieval from Lexical Memory: Roles in Inhibitionless Spreading Activation and Limited Capacity Attention*, 106 J. EXPERIMENTAL PSYCH. GEN. 226 (1977); Michael I. Posner & Charles R. Snyder, *Facilitation and Inhibition in the Processing of Signals*, in ATTENTION AND PERFORMANCE V 669, 669 (Patrick M.A. Rabbit & Stanislav Dornič, eds., 1975); David A. Taylor, *Time Course of Context Effects*, 106 J. EXPERIMENTAL PSYCH. GEN. 404 (1977). The general consensus in this literature is the delay can be attributed to both automatic and attentional components. In other words, some effects occur automatically as a result of processing the prime, and others involve the subject’s conscious expectations. Ira Fischler, *Associative Facilitation Without Expectancy in a Lexical Decision Task*, 3 EXPERIMENTAL PSYCH HUM. PERCEPTION AND PERFORMANCE 18 (1977); Proctor, *supra* note 62.

in an ad slows down processing because the consumer is busy wondering about what was behind the surprising stimulus. But because the small subject pool for our pilot response time studies increased the risk of spurious results, we decided to run our redesigned control condition with a large sample to confirm our suspicions. Accordingly, we designed a larger-scale experiment to more formally test the extent to which response time experiments measure dilution versus the effect of the unexpected. Since we found no indication that our context manipulation had any effect in either of our two pilots, we focused exclusively on the element of the unexpected.

### **C. Study 3: Response Time and Category Recall When Controlling For Surprise**

We proceeded with a larger scale study ( $N = 1343$ ) to disentangle the two possible causes for the observed reaction times in our two pilot studies. We randomly assigned subjects to a MERCEDES group or an INFINITI group. For each targeted mark, subjects were randomly assigned to one of three groups: a treatment group or one of two control groups. Subjects in the treatment group saw the three filler ads plus a MERCEDES or INFINITI diluting ad ( $N = 444$ ). Subjects in the first control group saw only the three filler ads ( $N = 451$ ) and subjects in the second control group saw the three filler ads plus a NIKE Toothpaste ad ( $N = 448$ ). After viewing the ads, subjects participated in the job search role play task, as a distractor. They then went on to answer the same 30 matching questions we presented in the first two studies.

#### **1. Response Times**

Our results indicate that once we control for the effect of any surprising mark-category combination, evidence of blurring disappears. In other words, with respect to the average response times with which subjects link MERCEDES and INFINITI with their traditional associations, there are no significant differences between subjects exposed to ads involving the targeted marks (MERCEDES toothpaste and INFINITI toothpaste) and subjects exposed to ads involving a mark other than the targeted marks (NIKE toothpaste).

Consider first Table 3A. This table reports the average treatment effects (i.e., the average response time of the treatment group minus the average response time of the control group) for MERCEDES treatment subjects (who were exposed to the three filler ads and the MERCEDES toothpaste ad) and control subjects within the first control group (who were exposed to the three filler ads only). In the first row, there is no statistically significant difference between the treatment and control groups on the product category pair MERCEDES-cars, which indicates that the diluting stimulus was not causing subjects to dissociate the brand MERCEDES from its traditional product

**Table 3A. Differences in Mean Response Times. Target Brand: Mercedes. No Control for Surprise.**

Treatment Group: Mercedes Toothpaste Ad				
Control Group: No 4thAd				
<b>Target Brand</b>				
<i>Mercedes-Luxury</i>	<i>Mercedes-Printers</i>	<i>Mercedes-Cars</i>	<i>Mercedes-Cheap</i>	<i>Mercedes-Wealth</i>
0.075**	0.023	-0.046	0.092	0.063
(0.041)	(0.053)	(0.105)	(0.079)	(0.066)
<b>Filler Brands</b>				
<i>Chase-Money</i>	<i>Chase-Restaurants</i>	<i>Coca Cola-Expensive</i>	<i>Coca Cola-Soft Drinks</i>	<i>Coca Cola-Sour</i>
0.025	0.037	-0.005	0.007	0.098**
(0.036)	(0.036)	(0.059)	(0.034)	(0.059)
<i>Coca Cola-Toothpaste</i>	<i>McDonald's-Affordable</i>	<i>McDonald's-Fries</i>	<i>McDonald's-Hamburgers</i>	<i>McDonald's-Healthy</i>
0.064	-0.008	0.067	0.060	0.042
(0.045)	(0.041)	(0.052)	(0.059)	(0.040)
<i>McDonald's-Sunset</i>	<i>Microsoft-Bicycles</i>	<i>Microsoft-Computers</i>	<i>Microsoft-Reliable</i>	<i>Microsoft-Software</i>
0.020	-0.038	0.164**	-0.012	0.024
(0.064)	(0.057)	(0.077)	(0.098)	(0.061)
<i>Microsoft-Luxury</i>	<i>Perrier-Bottle</i>	<i>Perrier-Clothing</i>	<i>Perrier-Pure</i>	<i>Perrier-Soft</i>
-0.021	-0.121	-0.071	0.041	-0.039
(0.058)	(0.105)	(0.067)	(0.088)	(0.077)
<i>Perrier-Water</i>	<i>United-Boxes</i>	<i>United-Horses</i>	<i>United-Small</i>	<i>United-Trucks</i>
0.019	-0.010	0.111	-0.077	0.080
(0.068)	(0.070)	(0.075)	(0.087)	(0.082)

Standard errors in parentheses.

One tailed test: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

category. We do see, however, what could be interpreted as a form of blurring of the association between MERCEDES and “luxury.” On average, subjects who saw the MERCEDES Toothpaste ad took 0.075 seconds longer than subjects in the control group to determine whether MERCEDES and LUXURY match ( $p < 0.05$ ). Although this difference is statistically significant, it is substantively very small. Furthermore, no other differences for the target MERCEDES brand are statistically significant from zero.

**Table 3B. Differences in Mean Response Times. Target Brand: Mercedes. Controlling for Surprise.**

<b>Treatment Group: Mercedes Toothpaste Ad</b>				
<b>Control Group: Nike Toothpaste Ad</b>				
<b>Target Brand</b>				
<i>Mercedes-Luxury</i>	<i>Mercedes-Printers</i>	<i>Mercedes-Cars</i>	<i>Mercedes-Cheap</i>	<i>Mercedes-Wealth</i>
-0.026 (0.096)	-0.096 (0.129)	0.022 (0.037)	0.047 (0.103)	0.083 (0.065)
<b>Filler Brands</b>				
<i>Chase-Money</i>	<i>Chase-Restaurants</i>	<i>Coca Cola-Expensive</i>	<i>Coca Cola-Soft Drinks</i>	<i>Coca Cola-Sour</i>
-0.014 (0.048)	0.029 (0.036)	-0.098 (0.073)	-0.066 (0.040)	0.032 (0.062)
<i>Coca Cola-Toothpaste</i>	<i>McDonald's-Affordable</i>	<i>McDonald's-Fries</i>	<i>McDonald's-Hamburgers</i>	<i>McDonald's-Healthy</i>
0.060 (0.042)	0.006 (0.036)	0.015 (0.073)	0.117** (0.050)	-0.017 (0.042)
<i>McDonald's-Sunset</i>	<i>Microsoft-Bicycles</i>	<i>Microsoft-Computers</i>	<i>Microsoft-Reliable</i>	<i>Microsoft-Software</i>
0.049 (0.056)	-0.079 (0.091)	0.104 (0.079)	-0.010 (0.060)	0.036 (0.050)
<i>Microsoft-Luxury</i>	<i>Perrier-Bottle</i>	<i>Perrier-Clothing</i>	<i>Perrier-Pure</i>	<i>Perrier-Soft</i>
-0.083 (0.058)	0.025 (0.050)	0.017 (0.046)	0.120 (0.077)	-0.033 (0.069)
<i>Perrier-Water</i>	<i>United-Boxes</i>	<i>United-Horses</i>	<i>United-Small</i>	<i>United-Trucks</i>
0.091 (0.065)	0.023 (0.062)	0.029 (0.080)	-0.012 (0.066)	0.021 (0.089)

Standard errors in parentheses.

One tailed test: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 3A further indicates that the diluting stimulus is associated with slower response times on two unrelated brands. Treatment group participants, for example, were slower by 0.098 seconds to react to the COCA-COLA-sour pair and by 0.164 to the MICROSOFT-computer pair. This signals that the

MERCEDES Toothpaste ad may have been unexpected by treatment subjects and slowed them down generally.

Table 3B shows what happens when we control for the appearance of a surprising mark-category combination. It reports the average treatment effects for MERCEDES treatment subjects and control subjects within the *second* control group (who were exposed to the three filler ads *and* the Nike Toothpaste ad). The statistically significant differences that we saw in Table 1A disappear, not simply with respect to MERCEDES-luxury but also with respect to COCA-COLA-sour and MICROSOFT-computers. Only one statistically significant response time difference remains (MCDONALD'S-hamburgers), which could be due to chance.

Tables 4A and 4B, which report results for the same experiments applied to the INFINITI mark, provide further support for the hypothesis that previous response time experiments measured the effect of having any unexpected stimulus. Table 4A reports treatment effects for the INFINITI treatment group as compared to the first control group, and Table 4B reports treatment effects for the INFINITI treatment group as compared to the second control group. Consistent with our MERCEDES results, Table 4B shows that all statistically significant differences disappeared when the second control group was also exposed to an ostensibly diluting ad.

On the basis of this and our two pilot studies, we conclude that previous response time studies used an insufficient control and thus cannot serve as models for surveys used in litigation or as evidence that ostensibly diluting conduct actually causes any appreciable blurring of consumers' ability to link marks with their traditional sources and associations. These previous studies used the equivalent of our first control group when they should have used the equivalent of our second control group. If future response time studies set out to detect dilution, whether in a litigation or academic context, they must expose control subjects to stimuli, that while not targeting the marks being tested for dilution, are at least as surprising or unexpected to control subjects as the ostensibly diluting stimuli to which treatment subjects are exposed.

## **2. Category Recall**

In addition to response time, Morrin and Jacoby examined the accuracy of category association matching as an indicator of trademark blurring. We did the same. We explored whether the MERCEDES Toothpaste and INFINITI Toothpaste ads had any impact on subjects' ability to recall the traditional product categories of the MERCEDES or INFINITI marks. If the MERCEDES Toothpaste or INFINITI Toothpaste ads are causing blurring, we would expect fewer subjects in the treatment groups to identify MERCEDES-cars or INFINITI-cars as a match compared to subjects in the control groups.

Table 5 shows average treatment effects for the category recall tests in terms of the difference in percentages of correct classifications between the treatment groups and the first and second control groups. The negative signs on

**Table 4A. Differences in Mean Response Times. Target Brand: Infiniti. No Control for Surprise.**

Treatment Group: Infiniti Toothpaste Ad				
Control Group: No 4thAd				
<b>Target Brand</b>				
<i>Infiniti-Luxury</i>	<i>Infiniti-Printers</i>	<i>Infiniti-Cars</i>	<i>Infiniti-Fast</i>	<i>Infiniti-Cheap</i>
0.029 (0.080)	-0.078 (0.078)	0.125 (0.096)	0.108 (0.110)	0.131** (0.053)
<b>Filler Brands</b>				
<i>Chase-Money</i>	<i>Chase-Restaurants</i>	<i>Coca Cola-Expensive</i>	<i>Coca Cola-Soft Drinks</i>	<i>Coca Cola-Sour</i>
-0.022 (0.033)	0.121** (0.066)	0.043 (0.045)	-0.023 (0.045)	0.024 (0.035)
<i>Coca Cola-Toothpaste</i>	<i>McDonald's-Affordable</i>	<i>McDonald's-Fries</i>	<i>McDonald's-Hamburgers</i>	<i>McDonald's-Healthy</i>
0.080 (0.117)	-0.038 (0.043)	-0.003 (0.038)	0.072** (0.036)	0.038 (0.038)
<i>McDonald's-Sunset</i>	<i>Microsoft-Bicycles</i>	<i>Microsoft-Computers</i>	<i>Microsoft-Reliable</i>	<i>Microsoft-Software</i>
0.120** (0.053)	-0.013 (0.061)	-0.012 (0.040)	-0.099 (0.084)	0.031 (0.039)
<i>Microsoft-Luxury</i>	<i>Perrier-Bottle</i>	<i>Perrier-Clothing</i>	<i>Perrier-Pure</i>	<i>Perrier-Soft</i>
0.065 (0.056)	0.039 (0.061)	0.231** (0.137)	-0.013 (0.067)	-0.024 (0.098)
<i>Perrier-Water</i>	<i>United-Boxes</i>	<i>United-Horses</i>	<i>United-Small</i>	<i>United-Trucks</i>
0.083** (0.035)	0.226 (0.161)	0.045 (0.075)	-0.073 (0.073)	0.043 (0.061)

Standard errors in parentheses.  
One tailed test: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4B. Differences in Mean Response Times. Target Brand: Infiniti.  
Controlling for Surprise.**

<b>Treatment Group: Infiniti Toothpaste Ad</b>				
<b>Control Group: Nike Toothpaste Ad</b>				
<b>Target Brand</b>				
<i>Infiniti-Luxury</i>	<i>Infiniti-Printers</i>	<i>Infiniti-Cars</i>	<i>Infiniti-Fast</i>	<i>Infiniti-Cheap</i>
0.021	-0.014	-0.043	0.079	-0.043
(0.054)	(0.055)	(0.141)	(0.101)	(0.107)
<b>Filler Brands</b>				
<i>Chase-Money</i>	<i>Chase-Restaurants</i>	<i>Coca Cola-Expensive</i>	<i>Coca Cola-Soft Drinks</i>	<i>Coca Cola-Sour</i>
-0.046	0.034	-0.068	-0.071	-0.045
(0.033)	(0.081)	(0.047)	(0.051)	(0.038)
<i>Coca Cola-Toothpaste</i>	<i>McDonald's-Affordable</i>	<i>McDonald's-Fries</i>	<i>McDonald's-Hamburgers</i>	<i>McDonald's-Healthy</i>
0.023	-0.080	-0.020	0.037	-0.045
(0.107)	(0.044)	(0.036)	(0.037)	(0.043)
<i>McDonald's-Sunset</i>	<i>Microsoft-Bicycles</i>	<i>Microsoft-Computers</i>	<i>Microsoft-Reliable</i>	<i>Microsoft-Software</i>
-0.045	-0.158	-0.047	-0.063	-0.060
(0.066)	(0.119)	(0.036)	(0.052)	(0.044)
<i>Microsoft-Luxury</i>	<i>Perrier-Bottle</i>	<i>Perrier-Clothing</i>	<i>Perrier-Pure</i>	<i>Perrier-Soft</i>
-0.032	0.002	0.216	-0.100	-0.094
(0.064)	(0.063)	(0.137)	(0.076)	(0.089)
<i>Perrier-Water</i>	<i>United-Boxes</i>	<i>United-Horses</i>	<i>United-Small</i>	<i>United-Trucks</i>
-0.018	0.081	-0.034	-0.110	-0.201
(0.041)	(0.172)	(0.076)	(0.082)	(0.121)

Standard errors in parentheses.

One tailed test: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 5. Category Recall Experiment. Difference between Treatment Group Mean and Control Group Mean.**

**A.**

---

**Treatment Group: Mercedes/Infiniti Toothpaste Ad**

**Control Group: No 4th Ad**

---

<i>Mercedes-Cars</i>	<i>Infiniti-Cars</i>
-0.006 (0.015)	-0.043* (0.027)

---

**B.**

---

**Treatment Group: Mercedes/Infiniti Toothpaste Ad**

**Control Group: Nike Ad**

---

<i>Mercedes-Cars</i>	<i>Infiniti-Cars</i>
-0.023** (0.012)	-0.016 (0.028)

Standard errors in parentheses.

One tailed test: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

the coefficients indicate that treatment group subjects were less likely to state that MERCEDES-cars or INFINITI-cars is a match. These differences, however, were generally quite small and statistically significant in just two out of four cases. Panel A shows comparisons between treatment subjects and subjects in the first control group (which were exposed to the three filler ads only). The difference between treatment and control subjects on the MERCEDES-CARS exercise was tiny and not statistically different from zero. In fact, 97.8 percent of subjects in the control group said these words were a match compared to 97.2 percent of subjects in the treatment group for a difference of 0.6 percent, or 0.006. The difference between treatment and control subjects on the INFINITI-CARS exercise is slightly larger and marginally statistically significant ( $p=0.055$ ). Control group subjects identified these words as a match in 93.3 percent of cases compared to treatment subjects, who identified the words as a match in 89 percent of cases. Turning to Panel B, however, the differences in the INFINITI-cars column do not hold up when the second control group saw a NIKE Toothpaste Ad. Control group subjects said INFINITI-cars is a match in 90.6 percent of cases compared to 89 percent of cases among treatment group subjects, a difference that is not statistically different from zero. In the MERCEDES-cars column in



Panel B, on the other hand, we now find that there is a statistically significant difference between treatment and control groups when the control group sees a NIKE Toothpaste ad. Control group subjects said these words were a match in 99.5 percent of cases compared to 97.2 percent of cases in the treatment group ( $p < 0.05$ ).

Taken together, the results presented in Table 5 show that subjects who saw a MERCEDES Toothpaste or INFINITI Toothpaste ad were always less accurate in tests that asked them to recall the primary product category of the MERCEDES or INFINITI marks. At the very least, this test indicates that seeing an ad for a familiar brand name in an unexpected product category can cause a small number of people to dissociate a brand name from its primary product category. This test cannot tell us, however, whether what we are observing is dilution or that a small number of subjects became confused or distracted after seeing the toothpaste ads—that is, they thought that the association we had actually presented to them (MERCEDES-toothpaste; INFINITI-toothpaste) was the product category association we were asking about. Recall that in Study 1, we found no evidence that exposure to an ostensibly diluting stimulus weakened the strength of association between the mark and its product category or relevant product attributes.<sup>65</sup> If the small reductions in product category recall we found for subjects in some, but not all, treatment branches in Study 3 were indeed related to product category dissociation caused by *blurring*, rather than by subject confusion, we would expect to have seen some reduction of the strength of associations in Study 1.

### III. Does Trademark Blurring Ever Occur?

The research reported here casts substantial doubt on the validity of the best evidence we currently have that the blurring form of trademark dilution actually exists. The response time studies that have been reported as (and assumed to be) evidence of blurring appear instead to be the result of treatment subjects' surprise at seeing ads associating famous marks with products very different from those normally sold under such marks. We applied a redesigned control condition that accounts for the surprise created by this unexpected stimulus, and the response delays that we were previously able to reproduce disappeared almost entirely.

Our experiments do not settle the question of the existence of trademark blurring. It is of course possible that future surveys used in litigation or academic empirical studies may be able to detect “association” between two similar marks that “impairs the distinctiveness” of one of them. But at present, we have no good evidence that this form of dilution actually occurs. Instead, we have little more than just-so stories, such as ours of a hypothetical Apple bar, that attempt to explain how dilution by blurring might operate.

So where does that leave us? In this part, we first consider an alternative interpretation of our results, one which we find unconvincing. We then suggest

---

<sup>65</sup> See Part I.B.2, *supra*.

that the theory of dilution by blurring remains such an unproven mess because, in practice, dilution is really about something else: misappropriation.

### **A. Superdilution?**

There may nevertheless be hope for the actuality of blurring in one interpretation of our results. This interpretation might hold that exposing consumers to surprising stimuli in the form of incongruous combinations of marks and new associations causes what might be termed “superdilution,” in which such surprising stimuli impair consumers’ ability to link marks with their traditional sources and associations for *all* marks throughout the marketplace.<sup>66</sup> Thus, when Brooklynites see Apple bar, their ability to navigate the marketplace as a whole marginally declines because they are reminded that the same mark may be used by more than one company.

We find this interpretation of our results unpersuasive. We believe that the increased response times we detected prior to the application of our redesigned control condition are an artifact of our experiment and not likely to translate to actual consumer perception in the marketplace. Subjects in our experiment are presented with a number of prompts, and they don’t know which prompts are salient. They are therefore likely, if they are engaged in the experiment, to treat every prompt as possibly salient. And therefore the surprising stimulus affects behavior in our experiment because it is treated by subjects as salient. Subjects do not know what the experiment is about. So they keep at the front of their mind all potentially relevant information. On the other hand, consumers in the marketplace rely on myriad contextual cues, as well as their own sense of their current aims, to decide which portion of the swarm of information coming at them every second to regard, and which portion to disregard. As a consequence, an unexpected stimulus in the marketplace may only rarely, or even never, be treated by subjects as salient. The surprising information is taken in as part of a welter of other information, and unless the environment is set up at that moment to make it salient, it may never be applied to a mark-product or mark-attribute identification effort. So, if one sees a “Mercedes Toothpaste” ad out in the marketplace, one might experience a moment of surprise, but for the stimulus to have an effect on behavior it would have to be salient to a particular task in which the observer is engaged or will soon become engaged. Otherwise, it seems unlikely that the surprise would affect behavior.

### **B. Dilution and Misappropriation**

Given the absence of proof that dilution by blurring actually occurs, we are left to speculate that the problem may ultimately not be with our experiments or the surveys or experiments of others, but with the concept of blurring itself. The notion of dilution originally emerged out of the 1924 German *Odol* case,<sup>67</sup> which

---

<sup>66</sup> This interpretation was offered by several legal scholars who commented on a prior draft of this article.

<sup>67</sup> LG Elberfeld, 14 September 1924, 204 – *Odol*; see generally Barton C. Beebe, *The Suppressed Misappropriation Origins of Trademark Antidilution Law: the Landgericht*

was essentially a case about misappropriation, one which happened to involve a trademark.<sup>68</sup> But when Frank Schechter introduced the concept of dilution to American law in his still widely-cited 1927 article *The Rational Basis of Trademark Protection*,<sup>69</sup> he downplayed the fact that dilution was a misappropriation doctrine and went so far as to expurgate from his translation of the *Odol* opinion the court’s core finding that the respondent had “appropriate[d] thus the fruits of another’s labor.”<sup>70</sup> Heavily influenced by law and economics thinking on trademark law, with its focus on consumer “search costs,” the American concept of dilution evolved into the concept of “blurring” and the idea that when two firms use the same mark, consumers must “think for a moment,” which increases their search costs.<sup>71</sup> We have struggled ever since to develop empirical proof that any appreciable increase in search costs actually occurs.

Yet courts continue to rule in favor of plaintiffs on antidilution causes of action. We suggest, as others have,<sup>72</sup> that they do so primarily for reasons sounding in misappropriation, but never articulated as such. The concept of “blurring” acts as an alibi for courts that want to reach what they deem to be the right result in cases, such as *Nikepal*, where a defendant uses another’s famous mark on unrelated goods. The defendant is not causing consumer confusion but it is reaping where it has not sown by appropriating to itself some of the notoriety of the famous mark. Courts are motivated to accept *mere association* survey evidence as full-blown evidence of impairment to justify enjoining conduct they deem to be unfair misappropriation.<sup>73</sup>

This is problematic for at least two reasons. First, we continue to twist ourselves into contortions to develop evidence of blurring, of some increase in response times and search costs, when courts’ underlying focus is on fairness and misappropriation. The result is that the concept of dilution grows increasingly vague and outcomes grow increasingly unpredictable. Second, unlike its

---

*Elberfeld’s Odol Opinion and Frank Schechter’s “The Rational Basis of Trademark Protection”*, in INTELLECTUAL PROPERTY AT THE EDGE: THE CONTESTED CONTOURS OF IP 59, 59 (Rochelle Dreyfuss & Jane Ginsburg eds., 2015).

<sup>68</sup> The *Odol* Court asserted that the damage to the mark was to its “selling power.” See Beebe, *supra* note 67.

<sup>69</sup> Frank I. Schechter, *The Rational Basis of Trademark Protection*, 40 HARV. L. REV. 813 (1927).

<sup>70</sup> See Beebe, *supra* note 67, at 72.

<sup>71</sup> See Posner, *supra* note 2.

<sup>72</sup> See, e.g., Apostolos Chronopoulos, *Goodwill Appropriation as a Distinct Theory of Trademark Liability: A Study on the Misappropriation Rationale in Trademark and Unfair Competition Law*, 22 TEX. INTEL. PROP. L.J. 253, 298 (2014); David Franklyn, *Debunking Dilution Doctrine: Toward a Coherent Theory of the Anti-Free-Rider Principle in American Trademark Law*, 56 HASTINGS L.J. 117 (2004); Richard A. Posner, *Misappropriation: A Dirge*, 40 HOUS. L. REV. 621, 623 (2003); David Welkowitz, *Reexamining Trademark Dilution*, 44 VAND. L. REV. 531 (1991).

<sup>73</sup> See *Nikepal*, 2007 WL 2782030, at \*8.

European counterpart,<sup>74</sup> the American antidilution statute contains no misappropriation provision. Congress has twice had the opportunity to include such a provision and twice it has declined to do so.<sup>75</sup> To the extent that courts are ruling on unstated misappropriation grounds, they are acting outside of the federal statutory framework.

#### IV. Conclusion: Future Directions for Research

The cause of action for trademark dilution by blurring has been authorized by federal law since 1995, and by the laws of various states for decades prior to that.<sup>76</sup> And yet the concept that the law identifies and prohibits may not even exist out in the world. Lawyers continue to litigate claims of dilution by blurring, but courts lack the most basic tool necessary for the disciplined adjudication of these claims – a reliable measure of the likelihood of dilution.

The *mere association* test is entirely unsuitable as a test for dilution by blurring. At the level of logic, evidence of the “association” between an ostensibly diluting use of a mark and the senior mark owner is not tantamount to evidence that the senior mark has become *disassociated* from the senior mark owner, as is necessary for the senior mark’s distinctiveness to be impaired. And as our Studies 1 through 3 demonstrate, an ostensibly diluting use may lead to association to a new category without causing the required disassociation from the existing one.

*Response time* tests may be a more promising measure of the likelihood of dilution, but they must be constructed carefully. We have attempted to do so here. In particular, we found that previous response time studies had failed to control for the surprising effect of the unexpected ostensibly diluting stimulus. Once we controlled for that surprise-type effect, thereby isolating in our experimental treatment the potential effect of dilution by blurring, we found that the evidence we’d previously found for blurring almost entirely disappeared.

This leaves us both better off and worse off than before. We believe that we have produced, at last, a methodologically sound test for measuring the likelihood of dilution. As applied to our test brands, however, the measure suggested that dilution did not occur. This is not to deny the possibility that the measure might yield evidence of dilution if applied to other target brands with other ostensibly diluting stimuli. That is possible. But even if we applied it to other brands and measured significant response delays indicating the possibility

---

<sup>74</sup> See Parliament and Council Directive 2015/2436, art. 10(2)(c), 2015 O.J. (L 336) 1, 11 (EC) (prohibiting the unauthorized use of a qualifying mark that “without due cause takes unfair advantage of, or is detrimental to, the distinctive character or the repute of the trade mark”).

<sup>75</sup> See Federal Trademark Dilution Act, Pub. L. No. 104-98, 109 Stat. 985 (1996) (codified at 15 U.S.C. § 1125(c)); Trademark Dilution Revision Act, Pub. L. 109-312, §120 Stat. 1730 (codified as 15 U.S.C. §1125(c) and replacing the Federal Trademark Dilution Act, Pub. L. 104-98).

<sup>76</sup> On state antidilution law, see MCCARTHY, *supra* note 5, at §24:77-83. By 2006, at least two-thirds of the states had some form of statutory antidilution law.

of dilution, additional questions would arise. We would need to test the possible mitigating effect of market context. And, most importantly, we would still need to test whether dilution by blurring that resulted in small time delays in the nature of milliseconds was a phenomenon likely to have real-world attribute association, preference, or purchase implications.

In the end, there are questions that remain open. But we have an intuition about where the evidence will lead. We suspect that the concept of dilution by blurring is a chimera—that is, an imaginary thing that we insist upon only because it is valuable to us as a proxy for attacking unauthorized uses of senior marks which our intuition tells us are unfair. It is true that in at least some dilution cases, defendants are engaged in some form of free-riding on the fame and goodwill enjoyed by the plaintiff's mark. Free-riding is, admittedly, not a simple thing to judge. Whether it is fair or unfair, productive or destructive, often depends on subtle differences in context. But courts are better at judging these sort of contextual questions than they are at running a social science inquiry into the hypothesized, but unproven, cognitive concept of "blurring." It is better, we think, to reframe the dilution cause of action away from cognitive science and toward the broader and more tractable questions of fairness and harm that the courts have become accustomed to dealing with in misappropriation cases.

## Appendix A: Amazon Mechanical Turk

We recruited subjects for our studies via Amazon's Mechanical Turk. Once a subject selected the task we listed on Mechanical Turk, he or she was directed to the Qualtrics online social science research platform to complete a survey instrument. Our instrument reflected the procedures used by Morrin and Jacoby and by Pullig, Simmons, and Netemeyer.

Online crowd-sourcing services like Amazon Mechanical Turk offer researchers the advantage of being able to run experimental protocols on large samples at a fraction of the cost of lab studies. They also allow researchers to pilot various instruments so that hypotheses and empirical tools can be tested, refined, and retested. Online response time experiments are possible because JavaScript timers can record clicks down to the millisecond. There are, however, some potential disadvantages to conducting a response time experiment online that we investigated and addressed. Namely, subjects in an online experiment use their own hardware and software, in contrast to a laboratory setting where subjects use the same hardware and software. While there are surely differences in the connection speeds, browser speeds, processor speeds, and keyboard and mouse clicking speeds of subjects recruited through Amazon Mechanical Turk, we assume that these differences are small and randomly distributed across subjects. Crump, McDonnell, and Gureckis confirm this assumption in a widely cited study that used subjects recruited through Mechanical Turk to replicate outcomes in four types of common laboratory reaction time studies.<sup>77</sup>

To test whether differences caused by hardware, software, and network speed are actually randomly distributed, we collected click-speed data from n=980 Mechanical Turk respondents using the timer function in Qualtrics. All respondents were required to use a desktop or laptop computer with mouse and asked to performed the same exact tasks. First, they were told to click the next button eleven times. Second, they were told to click a single option on the screen, then click the next button. They did this ten times. Third, they were asked five easy multiple choice questions and then told to click the next button (e.g., "Which company sells hamburgers? Microsoft, Google, McDonalds, or Facebook"). Qualtrics measured click speeds for each of these 26 tests to the thousandth of a second. We then randomly assigned respondents to two groups and compared mean click speeds for each of the 26 tests. If differences in hardware, software, and network speeds are not randomly distributed, we would expect to observe statistically significant differences between groups more than 5 percent of the time. We conducted 70 such randomized trials with group sizes ranging from n=50 to n=490. We found that differences between groups were statistically significant 4.77 percent of the time. This suggests that differences caused by hardware, software, and network speeds are randomly distributed across subjects.

---

<sup>77</sup> See Matthew J. C. Crump, John V. McDonnell & Todd M. Gureckis, *Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research*, 8 PLOS ONE (2013), <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0057410>.

Another potential disadvantage is that subjects recruited online cannot be monitored to ensure that they are following instructions or remaining attentive to the experiment. One instruction we feared might be broken was that subjects use a laptop or desktop computer with a mouse or touchpad, and not a touchscreen device like a tablet or smartphone. To ensure that we were only measuring reaction times of people using a mouse or touchpad, we applied a setting in Qualtrics that prohibited people with touchscreen mobile operating systems from participating in the experiment. To make sure subjects were paying attention to the experiment, we examined the data closely for invalid responses. We coded nine response times as missing values because they were greater than 50 seconds, which we believe signaled disengagement during the matching task. These excessively large response times constituted less than 0.02 percent of all response time values in our data set.<sup>78</sup> Ninety-nine percent of response times in our data set were 3 seconds or less, which we believe indicates that the vast majority of subjects were focused and engaged.

---

<sup>78</sup> During the data cleaning phase, we also found that a small fraction of responses times (1.08 percent) were zero. We coded these values as missing as well.

**Appendix B: Advertisements Used in Studies 1, 2, and 3**

Choose what's right for you

**CHASE BANK**

Chase Bank makes banking affordable and stress-free.

For all your banking needs

The bank you can rely on

CHASE BANK

*Note: Deposits at Chase Bank are federally insured.*

Expect to be moved

**UNITED VAN LINES**

United Van Lines movers are there when you need them

Best rates in the industry

From full-service to do-it-yourself moving & storage

UNITED VAN LINES

*Note: United Van Lines is not associated with United Airlines.*

Since 1863

**PERRIER SPARKLING WATER**

Perrier Sparkling Water is the best-tasting sparkling water

Natural, healthy, and pure

The ultimate refreshment to quench all thirsts

PERRIER SPARKLING WATER

*Note: Perrier Sparkling Water is regularly tested for purity.*



Toothpaste for the whole family

**INFINITI TOOTHPASTE**

Scientifically developed Infiniti toothpaste prevents decay with a fresh minty taste

Advanced whitening ingredients

Makes brushing your teeth a delight.

INFINITI TOOTHPASTE

*Note: Infiniti Toothpaste is not associated with Infiniti automobiles*

Toothpaste for the whole family

**MERCEDES TOOTHPASTE**

Scientifically developed Mercedes toothpaste prevents decay with a fresh minty taste

Advanced whitening ingredients

Makes brushing your teeth a delight

MERCEDES TOOTHPASTE

*Note: Mercedes Toothpaste is not associated with Mercedes Benz automobiles*