



I 
WAL-OCAUST
They have FAMILY VALUES
and their
ALCOHOL, TOBACCO
and FIREARMS
are 20% OFF

WAL-OCAUST
Come for the LOW prices
stay for the KNIFE fights
www.walo.caust.com

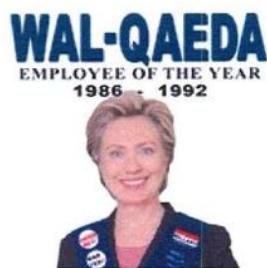
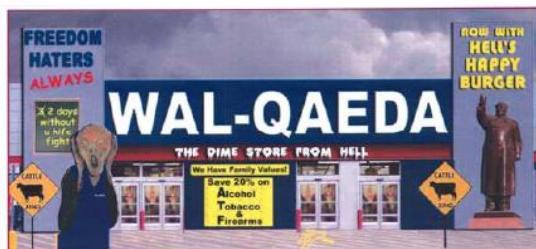
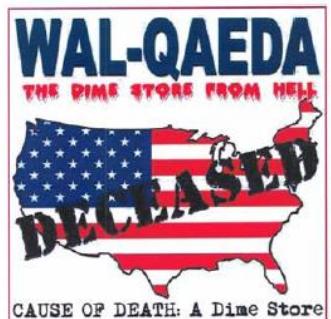


Appendix B: Challenged Wal-Qaeda Images

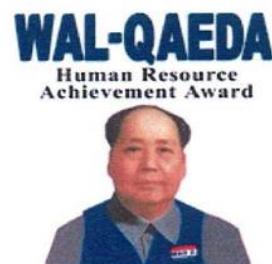


**FREEDOM
HATER
MART**

**SUPPORT
OUR
TROOPS
BOYCOTT
WAL-QAEDA**



**ATTENTION
WAL-QAEDA
THESE COLORS
DON'T RUN**



**BENTONVILLEBILIES
BENTONVILLEBULLIES**
ALWAYS

WAL-QAEDA
Freedom Haters **ALWAYS**

Smith v. Wal-Mart Stores, Inc.

537 F.Supp.2d 1302 (N.D.Ga. 2008)

Timothy C. Batten, Sr., District Judge:

....

II. Analysis

C. Trademark Infringement, Unfair Competition, Cybersquatting and Deceptive Trade Practices Claims

1. Actual Confusion

Proof of actual confusion is considered the best evidence of likelihood of confusion. *Roto-Rooter Corp. v. O'Neal*, 513 F.2d 44, 45–46 (5th Cir. 1975). A claimant may present anecdotal evidence of marketplace confusion, and surveys, when appropriately and accurately conducted and reported, are also widely and routinely accepted as probative of actual confusion. *See, e.g., AmBrit, Inc. v. Kraft, Inc.*, 812 F.2d 1531, 1544 (11th Cir. 1986) (considering the proffered survey but giving it little weight); *SunAmerica Corp. v. Sun Life Assurance Co. of Canada*, 890 F.Supp. 1559, 1576 (N.D.Ga. 1994) (viewing the proffered survey as confirmation of consistent anecdotal evidence).

Wal-Mart concedes that it has no marketplace evidence of actual consumer confusion. Instead, it presents two consumer research studies conducted by Dr. Jacob Jacoby that purport to prove that consumer confusion and damage to Wal-Mart's reputation are likely.

a. The Jacoby Report

Jacoby developed two surveys for Wal-Mart that both purported to measure consumer confusion and dilution by tarnishment. Specifically, the stated objectives of the research were (1) "To determine whether (and if so, to what extent), when confronted with merchandise bearing Mr. Smith's designs either in person or via the Internet, prospective consumers would be confused into believing that these items either came from Wal-Mart, came from a firm affiliated with Wal-Mart, or had been authorized by Wal-Mart," and (2) "To determine whether (and if so, to what extent) exposure to Mr. Smith's designs would generate dilution via tarnishment."

Deeming it impractical to test all of Smith's designs, Jacoby chose instead to test two products as representative of all of Smith's allegedly infringing products—the white t-shirt with the word "WAL*OCAUST" in blue font over the Nazi eagle

clutching a yellow smiley face, and another white t-shirt that depicted the word “WAL–QAEDA” in a blue font as part of the phrase “SUPPORT OUR TROOPS. BOYCOTT WAL–QAEDA.”

He also tested consumer reactions to “control” designs, which he compared to consumer responses to the Walocaust and Wal–Qaeda designs. To develop the control for the Walocaust design, Jacoby replaced the star with a hyphen and removed the smiley face from the yellow circle, and for both the Walocaust and Wal–Qaeda controls, he substituted “Z” for “W.” These substitutions resulted in control concepts entitled “Zal-ocaust” and “Zal–Qaeda.”

Jacoby engaged a market research firm to test each of the t-shirt designs in (1) a “product” study intended to test for post-purchase confusion and tarnishment, and (2) a “website” study intended to test for point-of-sale confusion and tarnishment.¹⁵

The market research company conducted the studies in a mall-intercept format. The company’s researchers would approach people who appeared to be thirteen years old or older and ask a series of screening questions.¹⁶ To qualify for either survey, the respondent was required to be at least thirteen years old¹⁷ and must have in the past year bought, or would in the coming year consider buying, bumper stickers, t-shirts or coffee mugs with words, symbols or designs on them. To qualify for the “website” study, the respondent must also have (1) used the Internet in the past month to search for information about products or services and (2) either (a) in the past year used the Internet to buy or to search for information about bumper stickers, t-shirts or coffee mugs with words, symbols or designs on them, or (b) in the coming year would consider buying over the Internet bumper stickers, t-shirts or coffee mugs with words, symbols or designs on them.¹⁸

¹⁵ This resulted in eight test cells:

	Test cells		Control cells	
Post-purchase confusion/tarnishment	Wal★ocaust t-shirt	Wal-Qaeda t-shirt	Zal-ocaust t-shirt	Zal-Qaeda t-shirt
Point-of-sale confusion/tarnishment	Wal★ocaust website	Wal-Qaeda website	Zal-ocaust website	Zal-Qaeda website

¹⁶ The research company conducted the surveys in malls in Trumbull, Connecticut; Philadelphia, Pennsylvania; Youngstown, Ohio; Chicago Ridge, Illinois; Louisville, Kentucky; San Antonio, Texas; Colorado Springs, Colorado; and Northridge, California. The website survey was also conducted in Portland, Oregon.

¹⁷ Because CafePress allowed only consumers over the age of thirteen to purchase from its site, Jacoby similarly limited his universe of respondents.

¹⁸ Respondents who worked at an advertising agency, a market research firm or a business located in the mall (or had an immediate family member who did) were

If the respondent met the qualifications, he or she was asked to go with the researcher to the mall's enclosed interviewing facility for a five-minute interview.¹⁹

For the "product" study, the interviewers presented to each respondent one of the four t-shirts described above and asked the respondent to imagine seeing someone wearing the shirt. The interviewer then asked a series of questions.

The first three sets of questions were designed to test for consumer confusion. The interviewers were directed to ask each of the "likelihood of confusion" questions sequentially unless the respondent answered "Sears," "Wal-Mart," "Youngblood's" or "K-Mart," in which case the interviewer was to record the answer, skip the remaining confusion questions, and go directly to the tarnishment questions.

In the consumer confusion series, the first set of questions tested for confusion as to source. The interviewer would ask "which company or store" the respondent thought "put out" the shirt, and if the respondent named a company or store, the interviewer then asked what about the shirt made the respondent think the shirt was "put out" by that company or store. The second set of questions, which dealt with confusion as to connection or relationship, asked the respondent whether the company or store that "put out" the shirt had some "business connection or relationship with another company" and if so, with what company. The respondent was then asked why he or she believed the companies had a business connection or relationship. A third set of questions, aimed at testing for confusion as to authorization or sponsorship, asked whether the company that "put out" the shirt needed permission from another company to do so, and if so, which company.

Finally, if the respondent had not yet answered "Sears," "Wal-Mart," "Youngblood's" or "K-Mart" to any of the first three sets of questions, he or she was then asked what the shirt made him or her "think of" and then "which company or store" the shirt brought to mind.

The fifth set of questions, which tested for dilution by tarnishment, were asked in reference to any company or store the respondent mentioned in his or her answers to the first four sets of questions. The first question asked whether seeing the shirt made the respondent more or less likely to shop at the store he or she had named, and the second question asked whether the perceived association with the store made the respondent more or less likely to buy the shirt.

excluded, as were people who normally wore eyeglasses or contact lenses but were not wearing them at the time of the screening.

¹⁹ The screening questionnaire provided to the Court indicates that the respondents who then participated in the surveys were given a monetary reward. Neither Jacoby's report nor any of the supporting survey documents disclosed the amount of the reward.

The interviews for the website study were much like those for the product study, except that instead of being shown the actual shirts, the respondents were exposed to a simulation of Smith's Walocaust CafePress homepage, his Wal–Qaeda CafePress homepage or the associated control homepage.²⁰ In each of the simulations, all of the hyperlinks were removed from the homepages except for the one hyperlink associated with the t-shirt that Jacoby had decided to test.

Jacoby directed the interviewers to begin each website interview by providing a URL to the respondent and asking the respondent to imagine that the URL was a search term the respondent had heard or seen somewhere and wanted to look up on the Internet. The interviewer would then have the respondent sit at a computer and type the URL into the browser. The URL would take the respondent to the simulated home page for testing.

The interviewer would then direct the respondent to look at the screen and scroll down the page “as [he or she] normally would” and click through to the first t-shirt on the screen. The respondent was then directed to click on the “view larger” box and look at the shirt as though he or she “found it interesting and [was] considering whether or not to order it” The interviewer would then ask the respondent exactly the same series of questions posed in the product study, including the same skip pattern to be applied in the event that the respondent mentioned Sears, Wal–Mart, Youngblood's or K–Mart in response to any of the consumer confusion questions.

In order to be tallied as “confused,” the respondent had to meet two tests. First, the respondent had to indicate either that the shirt came from Wal–Mart (first confusion series), came from a company that had some business connection or relationship with Wal–Mart (second confusion series), or came from a source that required or obtained permission from Wal–Mart (third confusion series). Second, the respondent had to indicate that his or her reason for that understanding was either because of the prefix “Wal,” the name (or equivalent), the smiley face, or the star after the prefix “Wal.” Thus, a respondent who believed that there was a connection between Wal–Mart and the t-shirt that he or she was shown but who did not mention the prefix “Wal,” the name (or equivalent), the smiley face, or the star, would not be counted as “confused.”

Any respondent who perceived an association between Wal–Mart and the t-shirt that he or she was shown and reported that the perceived association either made the respondent less likely to shop at Wal–Mart or more likely to buy that t-shirt was deemed to satisfy the requirement for dilution.

²⁰ The simulations were reproduced on a compact disc; the respondents did not view Smith's actual web pages on the Internet.

The field interviewers returned 322 completed interviews for the product study and 335 for the website study. Three responses were eliminated from the sample after the research company conducted a review to ensure that each respondent was qualified to participate in the study and that the questionnaires had been completed properly. The research company then sent the name and phone number of each of the interview respondents to an independent telephone interviewing service for validation, which consisted of calling each mall-intercept respondent to ensure that the respondent had actually participated in the study and that his or her answers were accurately recorded.

In the product study, 181 respondents (fifty-six percent of the usable sample) were positively validated, and sixteen respondents (about five percent) reported either different answers to the survey questions or claimed not to have participated in the study. The remainder either could not be reached during the twenty days Jacoby allocated for the validation or refused to respond to the validation survey.

Jacoby reported the results of those respondents who were positively validated plus the results from the respondents who could not be reached or would not respond to the validation survey, and he eliminated the results of the respondents who provided non-affirming answers during the validation process. This resulted in 305 reported responses to the product study: seventy-three for the Wal*ocaust concept, seventy-six for the Wal–Qaeda concept, seventy-nine for the Zal-ocaust concept, and seventy-seven for the Zal–Qaeda concept.

In the website study, 169 respondents (fifty-one percent of the usable sample) were positively validated, and forty-six respondents (about fourteen percent) reported either different answers to the survey questions or claimed not to have participated in the study. The remainder either could not be reached during the twenty days Jacoby allocated for the validation or refused to respond to the validation survey.

As he did in the product study, Jacoby reported the results of those respondents who were positively validated plus the results from the respondents who could not be reached or would not respond to the validation survey, and he eliminated the results of the respondents who provided non-affirming answers during the validation process. This resulted in 287 reported responses to the {website} study: seventy for the Wal*ocaust concept, seventy-eight for the Wal–Qaeda concept, sixty-nine for the Zal-ocaust concept, and seventy for the Zal–Qaeda concept.

Jacoby reported that the survey reflected high levels of consumer confusion and dilution by tarnishment. He claimed that the post-purchase confusion “product study” indicated a likelihood of confusion in nearly forty-eight percent of the respondents and that the point-of-sale confusion “website” study indicated a

likelihood of confusion in almost forty-one percent of the respondents.²¹ Jacoby also claimed that the “dilution” study indicated that almost twelve percent of the respondents were less likely to shop at Wal-Mart after seeing Smith’s designs.

b. Evidentiary Objections

Smith moves to exclude Wal-Mart’s expert report. He claims that Jacoby did not have the requisite Internet expertise to conduct the web-based “point-of-sale” portion of this particular study and that several aspects of Jacoby’s methodology affecting both portions of the study were faulty; thus, he contends, Jacoby’s study is “too deeply flawed to be considered . . .”

Wal-Mart argues that the Jacoby test was performed by a competent expert according to industry standards and therefore is valid. Wal-Mart further contends that the expert witnesses Smith presents in rebuttal are not experts in the area of consumer-goods “likelihood of confusion” trademark studies, and therefore *their* testimony is irrelevant and should be excluded.

Whether a given survey constitutes acceptable evidence depends on the survey’s ability to satisfy the demands of Federal Rule of Evidence 703, which requires consideration of the “validity of the techniques employed.” 233–34 FED. JUD. CTR., REFERENCE MANUAL ON SCI. EVIDENCE (2d ed. 2002) (explaining that in the context of surveys for litigation purposes, “[t]he inquiry under Rule 703[, which] focuses on whether facts or data are ‘of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject’ . . . becomes, ‘Was the . . . survey conducted in accordance with generally accepted survey principles, and were the results used in a statistically correct way?’”). See also *BFI Waste Sys. of N. Am. v. Dekalb County*, 303 F.Supp.2d 1335,1346 (N.D.Ga. 2004) (noting that the opposing party could have challenged an expert witness’s reference to a recent survey by questioning whether the survey methodology satisfied Rule 703).

The Eleventh Circuit has held that alleged technical deficiencies in a survey presented in a Lanham Act action affect the weight to be accorded to the survey and not its admissibility. *Jellibeans, Inc. v. Skating Clubs of Ga., Inc.*, 716 F.2d 833, 844 (11th Cir. 1983). Other courts have held that a significantly flawed survey may be excludable as evidence under either Rule 403 (the rule barring evidence that is more prejudicial than probative) or Rule 702 (the rule barring unreliable expert testimony). *Citizens Fin. Group, Inc. v. Citizens Nat'l Bank*, 383 F.3d 110, 188–21 (3d Cir. 2004) (finding that the district court properly excluded survey evidence under Rules 702 and 403 where the survey contained flaws that were not merely technical, but were so damaging to the reliability of the results as to be

²¹ Jacoby arrived at these numbers by averaging the net survey results for the Walocaust and Wal-Qaeda t-shirts.

“fatal”: the survey relied on an improper universe and its questions were imprecise); *Malletier v. Dooney & Bourke, Inc.*, 525 F.Supp.2d 558, 562–63 (S.D.N.Y. 2007). Even when a party presents an admissible survey purporting to show consumer confusion, however, the survey “does not itself create a triable issue of fact.” *Mattel, Inc. v. MCA Records, Inc.*, 28 F.Supp.2d 1120, 1133 (C.D.Cal. 1998) (citing *Universal City Studios, Inc. v. Nintendo Co.*, 746 F.2d 112, 118 (2d Cir. 1984), which found a survey “so badly flawed that it cannot be used to demonstrate the existence of a question of fact of the likelihood of consumer confusion”). . . .

To ground a survey as trustworthy, its proponent must establish foundation evidence showing that

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

Toys R Us, Inc. v. Canarsie Kiddie Shop, 559 F.Supp. 1189, 1205 (D.C.N.Y. 1983) (citing MANUAL FOR COMPLEX LITIG., 116 (5th ed. 1981), 4 LOISELL & MUELLER, FED. EVIDENCE § 472 (1979), and J. THOMAS MCCARTHY, TRADEMARKS & UNFAIR COMPETITION § 32:53 (1973)); accord *Rush Indus., Inc. v. Garnier LLC*, 496 F.Supp.2d 220, 227 (E.D.N.Y. 2007). Failure to satisfy any of the listed criteria may seriously compromise the survey’s impact on a court’s likelihood of confusion evaluation. *Id.*

Smith cites several grounds for excluding the Jacoby survey. He argues that the survey is inadmissible because it (1) failed to identify the relevant consumer universe or used a consumer universe that was substantially overbroad; (2) failed to replicate shopping conditions as consumers would encounter them in the marketplace; (3) was improperly leading; (4) violated the survey structure protocol necessary to comply with double-blind standards; and (5) failed to establish a relevant factual basis for Wal-Mart’s dilution by tarnishment claims. Smith further argues that even if the Court admits the survey, its consideration should be limited to only the two tested designs, despite Jacoby’s claim that they are representative of all the designs Wal-Mart seeks to enjoin.

As an initial matter, the Court observes that Smith does not take issue with Jacoby’s qualifications to design and conduct a consumer confusion survey and to analyze its results. It is undisputed that Jacoby is a nationally renowned trademark

survey expert who has testified hundreds of times. Smith contends, however, that Jacoby was unqualified to conduct this particular survey because he “lacks knowledge, experience, [and] sophistication” with regard to products marketed exclusively over the Internet and that as a result Jacoby’s survey protocol contained significant flaws.

Based upon its own review of Jacoby’s education and experience, the Court concludes that Jacoby is qualified to design and conduct a consumer survey and to testify about its results. To the extent that Jacoby’s purported lack of experience with surveys concerning goods sold exclusively online may have led him to test the wrong universe or to fail to replicate the shopping experience, as Smith has alleged, these factors will be examined when the Court evaluates the trustworthiness of the survey.

i. Web-Related Challenges

In undertaking to demonstrate likelihood of confusion in a trademark infringement case by use of survey evidence, the “appropriate universe should include a fair sampling of those purchasers most likely to partake of the alleged infringer’s goods or services.” *Amstar Corp. v. Domino’s Pizza, Inc.*, 615 F.2d 252, 264 (5th Cir. 1980). Selection of the proper universe is one of the most important factors in assessing the validity of a survey and the weight that it should receive because “the persons interviewed must adequately represent the opinions which are relevant to the litigation.” *Id.*

Similarly, “[a] survey that fails to adequately replicate market conditions is entitled to little weight, if any.” *Leelanau Wine Cellars, Ltd. v. Black & Red, Inc.*, 452 F.Supp.2d 772, 783 (W.D.Mich. 2006), *aff’d*, 502 F.3d 504 (6th Cir. 2007) (quoting *Wells Fargo & Co.*, 293 F.Supp.2d at 766). Although “[n]o survey model is suitable for every case . . . a survey to test likelihood of confusion must attempt to replicate the thought processes of consumers encountering the disputed mark or marks as they would in the marketplace.”

Smith hired Dr. Alan Jay Rosenblatt as a rebuttal witness to point out Internet-related deficiencies in Jacoby’s survey methodology—particularly deficiencies in universe selection and replication of marketplace conditions—that he claims resulted from Jacoby’s erroneous assumptions about how people reach and interact with websites.

. . . .

(a) Survey Universe

. . . . Wal-Mart maintains that Jacoby’s universe selection was proper. Smith counters that it was overly broad.

Although the universe Jacoby selected would include purchasers of Smith’s Walocaust or Wal-Qaeda merchandise, the Court finds that it is significantly overbroad. Because Smith’s merchandise was available only through his CafePress

webstores and the links to his CafePress webstores from his Walocaust and Wal-Qaeda websites, it is likely that only a small percentage of the consumers in the universe selected by Jacoby would be potential purchasers of Smith's products. A survey respondent who purchases bumper stickers, t-shirts or coffee mugs with words, symbols or designs on them may buy such merchandise because the imprint represents his or her school, company, favorite sports team, cartoon character, social group, or any of hundreds of other interests or affiliations; he or she may have no interest at all in purchasing merchandise containing messages about Wal-Mart, pro or con. The respondent may buy from brick-and-mortar stores or well-known retailers with Internet storefronts without being aware of Smith's website or CafePress, or may have little interest in buying such merchandise over the Internet at all. Therefore, a respondent who clearly falls within Jacoby's survey universe may nevertheless have no potential to purchase Smith's imprinted products. *See Leelanau Wine Cellars*, 452 F.Supp.2d at 782.

Other courts have similarly criticized surveys—including surveys Jacoby conducted in other trademark infringement cases—that failed to properly screen the universe to ensure that it was limited to respondents who were potential purchasers of the alleged infringer's product.

For example, in *Weight Watchers Int'l, Inc. v. Stouffer Corp.*, 744 F.Supp. 1259 (S.D.N.Y. 1990), Weight Watchers sued Stouffer for trademark infringement after Stouffer launched an advertising campaign that suggested that new exchange listings on Stouffer's Lean Cuisine packages would allow adherents to the Weight Watchers program to use Lean Cuisine entrees in their diets. *Id.* at 1262. Stouffer's likelihood of confusion survey, also conducted by Jacoby, identified the universe as "women between the ages of 18 and 55 who have purchased frozen food entrees in the past six months and who have tried to lose weight through diet and/or exercise in the past year." *Id.* at 1272. The court found that the universe was overbroad because the screener had not limited it to dieters, but also had included respondents who may have tried to lose weight by exercise only. The court concluded that as a result the survey likely included respondents who were not potential consumers, and because "[r]espondents who are not potential consumers may well be less likely to be aware of and to make relevant distinctions when reading ads than those who are potential consumers," that portion of the survey universe may have failed to make "crucial" distinctions in the likelihood of confusion testing. *Id.* at 1273.

....

(b) Shopping Experience

To be valid for the purposes of demonstrating actual confusion in a trademark infringement suit, it is necessary for a survey's protocol to take into account marketplace conditions and typical consumer behavior so that the survey may as

accurately as possible measure the relevant “thought processes of consumers encountering the disputed mark . . . as they would in the marketplace.” *Simon Prop. Group*, 104 F.Supp.2d at 1038; *accord WE Media, Inc. v. Gen. Elec. Co.*, 218 F.Supp.2d 463, 474 (S.D.N.Y. 2002).

Smith contends that Jacoby’s point-of-purchase study, which purported to measure consumer confusion over merchandise that Smith sold exclusively online, was improperly designed because it failed to take into account typical consumer Internet behavior. Wal-Mart does not contradict the expert testimony Smith proffers regarding consumer Internet behavior but instead maintains that it is irrelevant.

Jacoby’s point-of-purchase survey called for interviewers to provide each respondent with specific “search terms” that would take the respondent to a simulation of one of Smith’s websites. The respondent was asked to pretend that the resulting web page was of interest and to act accordingly (looking at the page and scrolling through it as the respondent would “normally” do), and then was directed to scroll down the page, below the first screen, and click on a specific t-shirt link. The respondent was not asked what message he or she took from the website or whether the website was in fact of interest. The survey protocol also gave the respondent no choice but to scroll down to the next screen and click on the t-shirt link, the only live link in the simulation.

In presenting Smith’s website and directing the survey respondents to click on one specific t-shirt link, Jacoby’s survey design presumed that all consumers who might be interested in a printed t-shirt, mug or bumper sticker would be equally likely to happen across Smith’s designs, regardless of the respondent’s level of interest in the messages on Smith’s webpage.

Although, as Wal-Mart points out, it is possible that some consumers may view web pages randomly and may scroll through and click on links on pages that are not of interest to them, the Court finds that the survey protocol did not sufficiently reflect actual marketplace conditions or typical consumer shopping behavior and therefore was unlikely to have elicited a shopping mindset that would have allowed Jacoby to accurately gauge actual consumer confusion.

Because Smith’s merchandise was available only through his CafePress webstores and the links to his CafePress webstores from his Walocaust and Wal-Qaeda websites, it is unlikely that many consumers randomly happen across Smith’s products. According to Rosenblatt’s uncontested testimony, people do not come to websites randomly, and they do not move within websites randomly. A great majority of Internet users arrive at a particular website after searching specific terms via an Internet search engine or by following links from another website. The user makes a judgment based on contextual cues—what is shown about a prospective website from the text of a search result or what is said about a

prospective website in the hyperlinked words and surrounding text of the website currently being viewed—in determining where to surf next. He moves from website to website, he moves within websites, and he performs actions such as signing a petition—or buying a product—by making choices based on what he sees and whether what he sees leads him to believe that going to the next page or following a link to another website will bring him to something he is interested in seeing, doing or buying.

In the marketplace, the visitor would be presented with a screen full of Smith's anti-Wal-Mart messages. Consumers who were interested in the messages on Smith's web pages would be motivated to choose the links that would eventually lead to his products, while those who were uninterested in Smith's messages would simply leave the page. Because the survey protocol directed the respondents to "pretend" to be interested in Smith's anti-Wal-Mart homepages and then directed them to click on a specific link, there is no assurance that the respondent actually read the homepage or would have been interested enough in it to be motivated to click on the t-shirt link. *See Gen. Motors Corp. v. Cadillac Marine & Boat Co.*, 226 F.Supp. 716, 737 (D.C.Mich. 1964) (observing that because survey respondents had little interest the allegedly infringing product, it followed that their inspection of the advertisement shown to them as part of the survey protocol was "casual, cursory and careless" and therefore of little probative value).

Other courts have similarly criticized surveys that failed to adequately replicate the shopping experience. In *Gen. Motors Corp.*, 226 F.Supp. at 737, the court criticized the proffered survey because it did not take into account typical consumer behavior:

Actual purchasers of a boat would not hastily read an advertisement, nor would a potential purchaser read it carelessly. A reasonable man, anticipating the purchase of a boat, would peruse the material at least well enough to note the manufacturer as being "Cadillac Marine & Boat Company, 406 Seventh Street, Cadillac, Michigan." Also, most buyers would want to see the boat itself before making a purchase.

Although the purchase of a t-shirt obviously does not involve the same level of financial consideration a consumer typically makes when buying a boat, a consumer is likely to consider the meaning of an imprinted t-shirt such as Smith's before wearing it in public. A reasonable person who was considering buying a t-shirt that references Al-Qaeda or the Holocaust would likely read the associated webpage at least well enough to see the harsh criticism of Wal-Mart and the prominent disclaimer dispelling any notion of a possible association with the company.

(c) Impact of Internet-Related Flaws on Survey's Evidentiary Value

For all of these reasons, the survey Jacoby conducted for Wal-Mart is of dubious value as proof of consumer confusion both because its survey universe was overinclusive and because its design failed to approximate real-world marketplace conditions. Jacoby's survey is subject to the same criticisms as his *Weight Watchers* survey . . . : Jacoby failed to screen the respondents to ensure that they would likely be aware of and make relevant distinctions concerning the specific product. *See Weight Watchers*, 744 F.Supp. at 1273. By failing to approximate actual market conditions, Jacoby further ensured that the survey would not "replicate the thought processes of [likely] consumers [of the junior user's merchandise] encountering the disputed mark . . . as they would in the marketplace." *See Simon Prop. Group*, 104 F.Supp.2d at 1038; *accord Gen. Motors Corp.*, 226 F.Supp. at 737. Therefore, the Court must consider these flaws in determining whether the survey is admissible and, if so, what evidentiary weight to afford it.

ii. Structural Flaws

Smith further alleges that the Jacoby study suffers from several structural flaws that diminish the trustworthiness of the results of both the web-based point-of-sale portion and the post-purchase t-shirt portion of the survey. He contends that (1) both the structure of the survey and the wording of several questions suggested the answers Wal-Mart wanted, and (2) the survey results should not be presumed to represent consumer reaction to any of the challenged merchandise that was not actually tested.

Smith hired Dr. Richard Teach as a rebuttal witness to point out deficiencies in Jacoby's website study survey methodology. Teach is an emeritus marketing professor and former dean at the Georgia Tech School of Business who has designed and conducted over one hundred surveys, including about fifty buyer surveys, and has taught survey methodology, statistics and related courses. Teach testifies that he agrees with Rosenblatt's testimony and also offers criticisms of his own. Smith uses Teach's survey expertise to support his *Daubert* argument that because the survey protocol contains multiple technical flaws, the results are unreliable and hence should be afforded very light evidentiary value if not completely excluded from evidence.

Wal-Mart moves to exclude Teach's testimony, supporting its motion with arguments much like those it used in its motion to exclude Rosenblatt's testimony

The Court finds . . . that his extensive experience designing and evaluating surveys qualifies him to provide testimony about technical flaws in the design of Jacoby's study and the impact of those flaws on the trustworthiness of Jacoby's reported results.

[T]o the extent that Teach's testimony focuses on general survey methodology, whether Jacoby's survey protocol deviated from standard methodology, and what impact any deviations may have had on the trustworthiness of Jacoby's reported results, Wal-Mart's motion to exclude it is DENIED.

(a) Leading Survey Structure and Questions

Smith argues that both the structure of the survey and the wording of several questions suggested the answers Wal-Mart wanted. Wal-Mart, of course, contends that Jacoby's survey presented no such risk.

(i) Double-Blind Survey Design

To ensure objectivity in the administration of the survey, it is standard practice to conduct survey interviews in such a way as to ensure that "both the interviewer and the respondent are blind to the sponsor of the survey and its purpose." REFERENCE MANUAL at 266. The parties agree that double-blind conditions are essential because if the respondents know what the interviewer wants, they may try to please the interviewer by giving the desired answer, and if the interviewer knows what his employer wants, he may consciously or unconsciously bias the survey through variations in the wording or the tone of his questions. *See id.*

Smith argues that the skip pattern included in Jacoby's survey hinted to the interviewers that Wal-Mart was the survey's sponsor. The survey protocol directed the interviewers to skip to the final tarnishment question, question five, if the respondent gave any one of four specific store names—Sears, Wal-Mart, K-Mart or Youngblood's—to any of the first three questions. Similarly, if the respondent did not give any of those four names in response to the first three questions, the interviewer was directed to ask "what other companies or stores" the stimulus t-shirt brought to mind, and only if the respondent answered with one of the four names was the interviewer to ask question five, the dilution question. The text on both of the tested t-shirts began with the prefix "Wal," and Wal-Mart was the only one of the four listed names that began with that prefix.

Smith argues that this series of questions combined with the t-shirt stimulus subtly informed the interviewers not only that a store name was desired, but also that a particular store name—Wal-Mart—was sought. Thus, Smith contends, because the survey failed to meet the double-blind requirement, it was not conducted in an objective manner and must be excluded for what must therefore be biased results. *See* REFERENCE MANUAL at 248 (noting that poorly formed questions may lead to distorted responses and increased error and therefore may be the basis for rejecting a survey).

Wal-Mart argues that the skip patterns followed proper protocol and that even if the interviewers guessed that Wal-Mart was involved, there could be no risk of bias because (1) interviewers are professionally trained and adhere to

extremely high ethical standards, and (2) it was impossible to determine from the design of the study who sponsored the study and for which side of a dispute the survey evidence was to be proffered.

Based on the facts that (1) both of the tested t-shirts include the prefix “Wal” and (2) the only store on the specified list of four that included that same prefix was Wal-Mart, it is safe to surmise that the interviewers at least suspected that Wal-Mart was involved in the survey in some manner. Aside from a common sense assumption that the party with deep pockets and reason to be insulted by the tested concepts was likely to have sponsored the research, however, the interviewers had no way to know who was the proponent of the research and who was the opponent. Thus, although the survey design may have breached generally accepted double-blind protocol to some degree, because the breach offered little risk of bias toward one party or the other the Court finds this issue to be of little import in its trustworthiness determination.

(ii) Leading Questions

Smith also argues that the wording of Jacoby’s confusion questions was improperly leading. Although the challenged t-shirts were created and offered for sale by Charles Smith, an individual, via his CafePress webstore, the survey asked about sponsorship only in the context of companies or stores, such as in the survey’s lead question, which asked, “[W]hich company or store do you think puts out this shirt?” Smith contends that this wording suggested to the respondent that the interviewer was looking for the name of a company or store, which would lead the respondent away from the answer that the shirt was put out by an individual who was criticizing a company. Wal-Mart counters that because Smith’s merchandise was sold through his CafePress webstores, the questions were accurately worded and thus not misleading.

The Court agrees with Smith that the disputed questions improperly led respondents to limit their answers to companies or stores. Though Smith did offer his merchandise through his CafePress webstore, as Wal-Mart argues, the Court finds this characterization disingenuous; the party Wal-Mart sued for offering the Walocaust and Wal-Qaeda merchandise for sale is not a company or a store, but instead Charles Smith, an individual. Furthermore, Wal-Mart has failed to point to any authority supporting the use of the “company or store” language in a consumer “likelihood of confusion” apparel survey or any such surveys previously conducted by Jacoby. Thus, the Court must consider this weakness in determining the admissibility or evidentiary weight to be accorded the survey.

(b) Representativeness

(i) Testing Stimuli

Smith also argues that the Jacoby survey results should not be presumed to represent consumer reaction to any of the challenged merchandise that was not

actually tested. Jacoby limited his surveys to testing two specific t-shirts (the Wal*ocaust smiley eagle shirt and the “SUPPORT OUR TROOPS” Wal–Qaeda shirt), and the conclusions stated in his report were narrowly drawn to refer to the tested t-shirts. At his deposition, however, he stated that because the tested shirts were “reasonably representative” of all the shirts that included the prefix “Wal” and the star, as in Wal*ocaust, or the prefix “Wal” and a hyphen, as in Wal–Qaeda, his results could be extrapolated from the tested t-shirts to all of the challenged t-shirts that shared those features.

Jacoby’s own deposition testimony supplies a fitting framework for analyzing this issue. When declining to offer an opinion about whether consumers would also be confused over the sponsorship of Smith’s Walocaust website, Jacoby stated that consumers respond differently to a given stimulus depending on the context in which it is presented, and because his survey tested only Smith’s CafePress webstores, his survey provided him with no data upon which to answer the question about consumer confusion regarding Smith’s website.

Applying the same reasoning, the Court finds that test results from one Walocaust or Wal–Qaeda t-shirt provide no data upon which to estimate consumer confusion regarding another Walocaust or Wal–Qaeda t-shirt. A consumer confused about the sponsorship of a shirt that says “SUPPORT OUR TROOPS [.] BOYCOTT WAL–QAEDA” may easily grasp the commentary in the more straightforwardly derogatory “WAL–QAEDA[.] Freedom Haters ALWAYS” concept. Similarly, a consumer confused over the sponsorship of a “Walocaust” shirt paired with an eagle and a smiley face might have a crystal clear understanding of the word’s meaning when it is superimposed over a drawing of a Wal–Mart–like building paired with a sign that advertises family values and discounted alcohol, firearms, and tobacco or when it is presented along with the additional text “The World is Our Labor Camp. Walmart Sucks.” As a result, this weakness will also impact the Court’s assessment of the survey’s evidentiary value.

(ii) Sample Size and Selection

Smith also challenges the survey’s small sample size; the Court additionally notes that Jacoby’s study employed mall-intercept methodology, which necessarily results in a non-random survey sample.

It is true that the majority of surveys presented for litigation purposes do, in fact, include small and non-random samples that are not projectible to the general population or susceptible to evaluations of statistical significance. 6 MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 32:165 (4th ed. 2006). Courts have found that “nonprobability ‘mall intercept’ surveys are sufficiently reliable to be admitted into evidence,” reasoning that because “nonprobability surveys are of a type often relied upon by marketing experts and social scientists in forming opinions on customer attitudes and perceptions,” they may be admitted

into evidence under Federal Rule of Evidence 703 as being “of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject.” *Id.*

However, probability surveys are preferred to non-probability surveys. *Id.* (citing Jacob Jacoby, *Survey & Field Experimental Evidence*, in SAUL KASSIN & LAWRENCE S. WRIGHTSMAN, JR., 185–86 THE PSYCHOLOGY OF EVIDENCE AND TRIAL PROCEDURE (1985)). Jacoby himself has written that “behavioral science treatises on research methodology are in general agreement that, all other things being equal, probability sampling is preferred to non-probability sampling.” Jacob Jacoby & Amy H. Handlin, *Non-Probability Sampling Designs for Litig. Surveys*, 81 TRADEMARK REP. 169, 170 (Mar.-Apr. 1991) (citing KUL B. RAI AND JOHN C. BLYDENBURGH, POL. SCI. STATS.. 99 (Holbrook Press Inc. 1973) and quoting its comment that “nonprobability samples do not represent the population truly, and the inapplicability of probability models as well as the impossibility of measuring or controlling random sampling error makes them even less attractive for scientific studies.”). Jacoby has similarly noted that although the vast majority of in-person surveys conducted for marketing purposes employ non-probability design, marketers more typically use telephone interviews, a “sizable proportion” of which employ probability designs. Jacoby & Handlin, 81 TRADEMARK REP. at 172 & Table 1 (estimating that sixty-nine percent of commercial marketing and advertising research is conducted by telephone).

Although courts typically admit nonprobability surveys into evidence, many recognize that “the results of a nonprobability survey cannot be statistically extrapolated to the entire universe,” and they consequently discount the evidentiary weight accorded to them. *Id.*; *accord Am. Home Prods. Corp. v. Barr Labs., Inc.*, 656 F.Supp. 1058, 1070 (D.N.J. 1987) (criticizing a Jacoby survey and noting, “While non-probability survey results may be admissible, they are weak evidence of behavior patterns in the test universe.”) Similarly, “[c]onducting a survey with a number of respondents too small to justify a reasonable extrapolation to the target group at large will lessen the weight of the survey.” 6 MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 32:171.

This Court finds troubling the Jacoby survey’s implicit assumption that a study protocol insufficient for many marketing purposes and heavily criticized for behavioral science purposes is nevertheless sufficient to aid a factfinder in a legal action challenging free speech. Therefore, this factor will also affect the Court’s assessment of the survey’s evidentiary value.

c. Admissibility

Having identified numerous substantial flaws in Jacoby’s survey, the Court must now determine whether the flaws limit the survey’s evidentiary weight or are so substantial as to render the survey irrelevant or unreliable and therefore

inadmissible under Federal Rule of Evidence 403, 702, or 703. *See Starter Corp. v. Converse, Inc.*, 170 F.3d 286, 297 (2d Cir. 1999) (excluding a survey under Rule 403 because the probative value of the survey was outweighed by potential prejudice and further noting that “a survey may be kept from the jury’s attention entirely by the trial judge if it is irrelevant to the issues”) (citing *C.A. May Marine Supply Co. v. Brunswick Corp.*, 649 F.2d 1049 (5th Cir. 1981)); *accord Ramdass v. Angelone*, 530 U.S. 156, 173, 120 S.Ct. 2113, 147 L.Ed.2d 125 (2000) (listing numerous cases in which courts have excluded or minimized survey evidence as unreliable).

Courts in the Eleventh Circuit typically decline to exclude likelihood of confusion surveys and instead consider a survey’s technical flaws when determining the amount of evidentiary weight to accord the survey. *See, e.g., Jellibeans*, 716 F.2d at 845; *Nightlight Sys., Inc. v. Nitelites Franchise Sys., Inc.*, 2007 WL 4563873 at *5 (N.D.Ga. Jul.17, 2007). Consequently, although this is a close case, the Court concludes that the better option is to admit the survey evidence and to consider the survey’s flaws in determining the evidentiary weight to assign the survey in the likelihood of confusion analysis.

The Court finds, however, that because the survey tested only the “SUPPORT OUR TROOPS[.] BOYCOTT WAL–QAEDA” t-shirt and the Walocaust eagle t-shirt, it has no relevance to any of Smith’s other Wal–Mart–related concepts. The Court agrees with Jacoby that context matters—a lot—and therefore will not consider Jacoby’s survey as evidence of likelihood of confusion with regard to the words “Walocaust” and “Wal–Qaeda” in general; the study is admissible only as to the two concepts that Jacoby actually tested. *See Fed.R.Evid. 702* (limiting expert testimony to that “based upon sufficient facts or data”).

Even with regard to the tested concepts, the Court finds that the survey was so flawed that it does not create a genuine issue of material fact. *See Spraying Sys. Co. v. Delavan, Inc.*, 975 F.2d 387, 394 (7th Cir. 1992) (recognizing that if a proffered survey is severely and materially flawed, it may not be sufficient to establish a genuine issue of material fact even if it purports to show evidence of actual confusion). Jacoby surveyed an overbroad universe, failed to adequately replicate the shopping experience, and asked leading questions. He also surveyed a non-random sample that in any case was too small to allow the results to be projected upon the general market. Thus, the Court finds that the Jacoby survey is so flawed that it does not establish a genuine issue of material fact with regard to actual confusion, much less *prove* actual confusion.

Lack of survey evidence showing consumer confusion is not dispositive, however; the Eleventh Circuit has moved away from relying on survey evidence. *Frehling Enters. v. Int’l Select Group, Inc.*, 192 F.3d 1330, 1341 n. 5 (11th Cir. 1999). In fact, a court may find a likelihood of confusion in the absence of any evidence of actual confusion, even though actual confusion is the best evidence of likelihood of

confusion. *E. Remy Martin & Co. v. Shaw-Ross Int'l Imps., Inc.*, 756 F.2d 1525, 1529 (11th Cir. 1985). Accordingly, the Court will now consider the remaining likelihood of confusion factors.

{The court ultimately granted summary judgment to Smith on Walmart's confusion and dilution claims.}