An Examination of the Psychometrical Comparability of Survey Evidence in Sport Trademark Litigation

Sungho Cho and Anita M. Moorman

Parties in trademark infringement litigation frequently introduce survey evidence to demonstrate consumer confusion. Especially, *Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Ltd. Partnership* (1994) highlights the significant roles of such survey evidence in sport trademark litigation. While *Colts* is arguably one of the most influential federal cases in sport trademark jurisprudence, the decision has been criticized for various grounds. This article focuses on the questionable validity of the survey evidence used in *Colts*. After discussing the crucial values of survey information in trademark infringement cases, it points out the problematic psychometrical validity of the key evidence used by the plaintiffs in *Colts*. Lastly, the paper introduces an *SEM* protocol widely used in social science research (i.e., Congenerity Test; Ohanian, 1990), as a gatekeeping technique that would likely screen out unreliable survey information. It demonstrates how the technique could have been used by the parties in *Colts* case.

This article examines the use of multivariate statistical analysis in sport trademark infringement cases. Survey evidence has been regularly incorporated into trademark litigation to prove consumer confusion which is an essential element of a trademark infringement legal claim. However, the use of survey evidence in trademark litigation has been subject to vague and inconsistent standards (Manta, 2007). These vague and inconsistent standards for using survey evidence has produced a continuing and compelling challenge for trademark holders to successfully pursue trademark infringement claims. In addition, the lack of consistent standards for survey evidence also prevents companies who use marks similar to protected trademarks from demonstrating that their use is not in violation of trademark laws. Thus, all potential parties to trademark litigation would benefit from more predictable and consistent use of survey evidence in trademark litigation. Survey evidence has been used in a number of sports trademark infringement cases, most notably in *Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Ltd. Partnership* (1994; hereinafter "*Colts*"). A central issue in the *Colts* case was the appropriate and permissible use of survey evidence. While survey evidence has been used in many sport trademark cases, the admissibility and validity of the evidence was not a central issue as it was in the *Colts* decision. Thus, this decision presents a

Cho is with the School of HMSLS, Bowling Green State University, Bowling Green, OH. Moorman is with the Dept. of Health & Sport Sciences, University of Louisville, Louisville, KY.

unique real world, rather than hypothetical, opportunity to examine and illustrate the impact of using multivariate testing to validate survey evidence. This article examines how survey evidence was used in the *Colts* case to determine whether a structural analysis such as Congenerity Test (Ohanian, 1990) provides a more scientifically grounded testing procedure for the validity and accuracy of consumer confusion survey evidence.

First, the article provides an introduction to the jurisprudence of trademark law particularly related to sport properties or organizations and the use of survey evidence in trademark litigation. Second, the article discusses how the Congenerity Test would have provided the defendant in the *Colts* case with a stronger scientific foundation to challenge the plaintiffs' survey evidence. Finally, this article recommends structural equation modeling ("SEM" hereafter) as a preferred statistical approach for validating consumer confusion survey evidence in trademark litigation.

Use of Survey Evidence in Trademark Litigation

At the core, a trademark is a product or service source identifier. Trademarks serve to identify the source or origin of a product or service so that consumers can make informed purchase decisions and locate information in the marketplace more easily. One underlying public policy traditionally supporting federal trademark laws is the protection of consumers from being deceived or confused as to source or origin of goods or services (Aaker, 1996; Grady & McKelvey, 2008). Thus, a recurring question in trademark infringement actions is whether an appreciable number of consumers are likely to be confused about the product information provided. (Geise, 1995). This element of a trademark infringement action is known as "likelihood of consumer confusion" and proof of actual consumer confusion is an important factor in the likelihood of confusion analysis. (Geise, 1995). To prove actual consumer confusion, survey evidence has become widely used and accepted in trademark infringement litigation. But such evidence can be met with skepticism in some courts, and often times litigants will offer dueling expert testimony based on survey evidence. For survey evidence to produce meaningful evidence of consumer confusion, the state of mind of a consumer needs to be directly measured or extrapolated using preferred research techniques that can be clearly and easily understood by a judge or jury. This section will provide a basic overview of trademark law and trademark infringement litigation, and discuss several cases illustrating challenges faced by litigants using survey evidence to demonstrate consumer confusion.

Overview of Federal Trademark Law and Trademark Infringement

The Federal Trademark Act of 1946 (hereinafter the "Lanham Act") prohibits the unauthorized use of a registered trademark. The Lanham Act permits trademark owners to bring a cause of action against any person using protected marks without the trademark owner's permission. This unauthorized use is known as "infringement" and section 1114 of the Lanham Act, permits an owner of a federally registered trademark to bring a trademark infringement action when another person

use[s] in commerce any reproduction, counterfeit, copy or colorable imitation of a registered mark in conjunction with the sale, offering for sale, distribution or advertising of any goods or services on or in connection with which such use is *likely to cause confusion* or to cause mistake, or to deceive (15 U.S.C. § 1114(1)(a), 2005) (emphasis added).

In addition to prohibiting trademark infringement, the Lanham Act also prohibits trademark dilution, false designation of the origin of a product/service, and unfair competition (15 U.S.C. §1125, 2005). Section 1125 provides

Any person who, on or in connection with any goods or services, or any container for goods, uses in commerce any word, term, name, symbol, or device, or any combination thereof, or any false designation of origin, false or misleading description of fact, or false or misleading representation of fact, which

(A) is *likely to cause confusion*, or to cause mistake, or to deceive as to the affiliation, connection, or association of such person with another person, or as to the origin, sponsorship, or approval of his or her goods, services, or commercial activities by another person, or

(B) in commercial advertising or promotion, misrepresents the nature, characteristics, qualities, or geographical origin of his or her or another person's goods, services, or commercial activities,

shall be liable in a civil action by any person who believes that he or she is or is likely to be damaged by such act (15 U.S.C. § 1125(a)(1)(2005) (emphasis added).

The protections provided in section 1125 are much broader than the trademark infringement protections contained in Section 1114 (Grady & McKelvey, 2008). However, trademark infringement, trademark dilution, and unfair competition claims all require proof of consumer confusion. The role of consumer confusion to Lanham Act claims is often debated among scholars who examine the origin and evolution of the Lanham Act since its adoption. For example, trademark law has been said to be rooted in consumer protection principles, and more precisely designed to improve quality of information in the marketplace and reduce consumer search costs (McKenna, 2007). Trademarks aid consumers identification and organization of product information which prevents consumer deception and thereby reduce costs of searching in the marketplace (McKenna). Lemley (1999) argued that the Lanham Act has never intended to maximize profits for trademark owners at the expense of competitors and general consumers. In fact, the legislative history of Lanham Act indicates that Congress did not intend to create unfettered exclusive property rights for trademark owners except to the extent that might be necessary to prevent unauthorized or deceptive use of a protected mark (Kahn, 2004). Trademarks function to facilitate the repeated purchases of goods and services that have delivered positive prior consumption experience to consumers by simplifying their information search process.

But more recent doctrinal innovations, such as trademark dilution, trade dress protections, trademarks as domain names, and licensing practices, seem to migrate away from traditional trademark principles and reflect a property based principle of

trademarks. Some scholars suggest that traditional trademark law sought to protect producers from competitors' illegitimate diversions of trade, more than it sought to protect consumers (McKenna, 2007). Thus, proof of consumer confusion was simply necessary to assess the competitor's actions as legitimate or illegitimate. Regardless whether trademark law is designed to protect consumers, improve quality of information in the marketplace, or establish competitive property rights, some amount of consumer confusion must be present for a trademark owner to prevail in trademark infringement case (Kahn, 2004).

According to Brogan (1996), likelihood of confusion is defined as the probability that the accused infringer's mark is the legal cause in fact, of confused, mistaken or deceived states of mind of potential consumers. A plaintiff has the burden to prove the likelihood of confusion in cases of trademark infringement (Brogan, 1996). The doctrine of likelihood of confusion in trademark law accommodates a legal notion that the schematic association between a trademark and goods or services designated by the mark is the central value of the mark. This psychological association is presumed to facilitate consumers' repeated purchases of some goods or services based on their positive consumption experience and/or in response to other marketing efforts conducted by the mark's owner. Therefore, if this association is somewhat disrupted by another deceptively similar mark or other allegedly infringing activities, it would devalue the core function of the trademark and consumer confusion would likely ensue (Manta, 2007).

To examine the likelihood of confusion in trademark infringement claims, federal circuits have developed unique multifactor tests (Manta, 2007). One of the most widely known versions of the test is the multiple factor test developed by the Second Circuit in *Polaroid Corp. v. Polarad Electronics Corp*. (1961), which consists of eight factors: (1) strength of plaintiff's mark; (2) the degree of similarity between the two marks at issue; (3) the proximity of the products; (4) the likelihood that the prior owner will bridge the gap; (5) actual confusion; (6) the reciprocal of defendant's good faith in adopting its own mark; (7) the quality of defendant's product; and (8) the sophistication of the buyers. In general, no single factor will control the likelihood of confusion analysis. However, prong 5, actual confusion, is considered one of the most influential factors because it can decisively bolster the entire rationale of trademark infringement based on some tangible evidence (Thornburg, 2005). While a variety of evidence could help to demonstrate actual confusion, survey evidence is particularly helpful to establish consumer confusion on a broad scale.

Survey Evidence in Trademark Infringement Cases

It is theoretically desirable that all consumers who are likely confused by the defendant's action physically appear in the courtroom and testify about their perception of the allegedly infringing practice relating to the plaintiff's mark. Courts may not, however, have such capacity or be willing to perform the logistically impossible task. Therefore, trademark law allows parties to introduce some social science information presenting the state of mind of consumers in trademark litigation such as anecdotal evidence (Wheble, 2009), consumer survey evidence (Jones, 1989), and experimental data (Lipton, 1987). Among these, a consumer survey might be the most widely used type of social science methodology in the field because it

can measure the state of mind of a consumer more efficiently with relatively less cost. Moreover, the reliability of the information can be systematically scrutinized by looking into the protocols of the data collection and analysis (Lipton, 1987).

In *Zippo Manufacturing Co. v. Rogers Imports, Inc*. (1963), the court announced that survey evidence is admissible under at least two grounds with respect to hearsay rule. First, survey information may come under the state of mind exception to the hearsay rule. Secondly, the balance between the need for survey evidence and the circumstantial guarantees of trustworthiness may warrant the admission of the type of information especially where the psychological state of mind of consumers is at stake. While *Zippo* held that survey information is admissible in trademark infringement cases, it also accentuated that the admissibility of the survey evidence still depends on the "circumstantial guarantees of trustworthiness" (*Zippo Manufacturing Co. v. Rogers Imports, Inc*., 1963) that must be proved by the party who introduces survey evidence (Brogan, 1996). McCarthy (2011) enumerates eight factors to decide the reliability of survey evidence: (1) whether the proper universe was examined; (2) whether a representative sample was drawn from the universe; (3) whether the mode of questioning the interviewee was proper; (4) whether the persons conducting the survey are recognized experts; (5) whether the data gathered was accurately reported; (6) whether the sample design was correct; (7) whether the actual questionnaire given to interviewees was not leading; and (8) whether the overall interviews were performed in accordance with objective statistics in the applicable field.

Survey information is also admissible under Rule 703 and 702 of the Federal Rules of Evidence. Under the statute, it can be introduced in conjunction with an expert opinion on the ground that an expert is allowed to base his or her opinion on survey data as far as the methodology is reasonably relied upon in the expert's particular field in forming opinions or making inferences (Jones, 1989). The use of expert testimony is to assist the trier of fact (Notes of advisory committee, FRE, 702). The admissibility of expert testimony essentially rests on whether the testimony will aid the trier of fact and is reliable based on available methodology. An expert would likely testify concerning both the validity of the methodology for the survey and the meaning of the results. Survey evidence in trademark litigation would be also admissible in terms of the federal standard for scientific evidence articulated in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993). In *Daubert*, the U.S. Supreme Court emphasized the "gatekeeper" role of judges in dealing with scientific evidence and enumerated four factors that might be used as criteria in the consideration of the reliability of all scientific evidence: (1) whether a technique at issue has been tested; (2) whether a technique has undergone peer review; (3) whether a technique provides its error rate; and, (4) whether a technique has been accepted by the relevant scientific community (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 1993). The *Daubert* standard has also been extended to nonscientific expert testimony in *Kumho Tire Co. v. Carmichael* (1999). Since consumer surveys are without a doubt a prevalent method in social science research relating to data collection, statistical analysis, manipulation check, etc., most consumer survey evidence would certainly satisfy *Daubert*. It is generally acknowledged that the questionable reliability of survey information in light of the McCarthy's multifactor test would likely affect the weight of the evidence rather than the admissibility (Sarel & Marmorstein, 2002; Thornburg, 2005). In some cases, the absence of

survey evidence might be fatal, especially when a plaintiff has the financial capacity to conduct a survey (*Medici Classics Prods, LLC. v. Medici Group, LLC.*, 2008; *Playboy Enterprises, Inc. v. Netscape Communications Corp.*, 1999; *Eagle Snacks, Inc. v. Nabisco Brands, Inc*., 1985).

Survey Evidence in Sport Trademark Litigation

Given the significant commercial value directly and indirectly associated with names and logos, sport organizations are implementing far more aggressive tactics to protect their trademark rights (Hylton, 2011). A number of trademark disputes in the sport industry have been litigated on the issue of survey evidence and consumer confusion. For example, in 1991, Major League Baseball ("MLB" hereafter) granted a new team franchise in Colorado. Although the team initially considered "Bears," it finally chose "Rockies" anticipating possible trademark disputes with the Chicago Bears. In fact, "Rockies" was the name of an old National Hockey League ("NHL" hereafter) franchise in Colorado which had moved to New Jersey in 1984. While the NHL could have defended the trademark rights associated with "Rockies," it did not challenge the new baseball franchise's name selection (Mark & Jacoby, 2006). The generosity of the NHL Rockies has rapidly disappeared in the industry ever since (Angel, 2010). In addition, when Cleveland Browns relocated the National Football League (hereafter "NFL") franchise to Baltimore in 1996, Clevelanders vigorously fought to keep the franchise and name (Babington & Denlinger, 1995). Although they failed to keep the franchise in the city, a series of lawsuits by fans, ticket holders eventually lead to a settlement that allowed city to retain the team's trademark rights and name (*See*, Leone, 1997). When the new Cleveland Browns resumed its NFL season in 1999, the trademark was reinstated as being used in commerce again (Angel, 2010). In another case, when the Montreal Expos moved to Washington D.C. in 2006, it found that a Cincinnati apparel company, Bygone Sports, had already registered the old local MLB franchise, Washington Nationals, which was the desired name for the new team (Heath, 2006). MLB Properties and the Expos filed a lawsuit against Bygone Sports to reclaim the "Nationals" name (*Major League Baseball Properties v. Bygone Sport*s, 2005). Eventually, the parties reached a settlement but these examples serve to illustrate the recurring costs and significance of trademark protection in the sport industry ("Nationals settle trademark lawsuit," 2006; Angel, 2010).

In addition to the examples provided above, two well-known trademark infringement cases have shaped sport trademark protection policies for many years. Both also involved the use of survey evidence to prove consumer confusion. First, In *MLB Properties, Inc. v. Sed Non Olet Denarius, Ltd*. (1993) the Los Angeles Dodgers and MLB Properties sued the Brooklyn Dodger Sports Bar and Restaurant in New York City for trademark infringement. The federal district court held that MLB was unable to prove actual or a likelihood of confusion specifically rejecting their survey evidence as fatally flawed or irrelevant to the issue of actual confusion. With regard to the survey evidence the court stated that the "survey evidence which, if credible and not fatally flawed by its structure or content, might have been probative on the actual confusion factor." (*MLB Properties, Inc. v. Sed Non Olet Denarius, Ltd.*, 1993, p. 1119). The survey evidence was rejected because they only demonstrated some associational connection rather than consumer confusion

regarding ownership or authorization. The court found that several questions in the survey were suggestive or leading resulting in invalid responses. Therefore, the court concludes that there is no proof of actual confusion. The court also determined in favor of the defendants on their affirmative defense of abandonment concluding the baseball plaintiffs' had abandoned the Brooklyn Dodgers trademark when the team left New York and changed the franchise name to "Los Angeles Dodgers." According to the court, the word "Brooklyn Dodgers" was not just a name of baseball organization with the geographical designation but a "non-transportable cultural institution" independent from the baseball team which had been existing in the public domain since the team moved to California decades ago (*MLB Properties, Inc. v. Sed Non Olet Denarius, Ltd.*, 1993, p. 1128; Brogan, 1996). *MLB Properties* was not a binding precedent in the next significant case *Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Ltd. Partnership* (1994) but does bear some striking similarities in that a former team name was being used and survey evidence was offered to prove consumer confusion.

In 1984, Baltimore Colts, which had a rich history in Maryland, suddenly moved the franchise to Indianapolis. This abrupt and secretive relocation of the team infuriated Baltimoreans. The city even attempted to get the franchise back by exercising its eminent domain power but eventually failed (*Indiana Colts v. Mayor & City Council of Baltimore*, 1984). Nine years later, when the Canadian Football League ("CFL" hereafter) granted a new CFL franchise for the city of Baltimore, the new team named itself "Baltimore CFL Colts" in response to the public outcry for the old team name. The NFL and Indianapolis Colts immediately brought an action against Metropolitan Baltimore Football Club (hereinafter "Baltimore") for trademark infringement under Lanham Act 43(a) to enjoin the CFL team from using the name "Baltimore Colts."

The federal district court decided in favor of the NFL on the grounds that even though neither the Indianapolis Colts nor NFL renewed the federal registration of the old "Baltimore Colts" mark, "Baltimore CFL Colts" would likely confuse a significant number of ordinary consumers in the relevant market (*Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Ltd. Partnership*, 19944, p. 416). That is, the court concluded that consumers may mistakenly believe that the new CFL team was somehow affiliated with the Indianapolis Colts or the NFL. Baltimore appealed to the federal court of appeals where the appellate court affirmed the trial court's decision. Unlike the MLB in *Sed Non Olet*, the NFL successfully used consumer surveys to demonstrate that the source of consumer confusion was not the word "Baltimore" but instead the word "Colts," presumably to avoid non-transferable cultural institution arguments which could have been asserted by the defendant. The appellate court pointed out that Baltimore CFL Colts failed to rebut the NFL's survey evidence showing the likelihood of confusion associated with the "Baltimore CFL Colts" name. The *Colts* decision solidified trademark rights of relocating sport franchises. The case effectively held that a team may enforce trademark rights associated with its old name even after it completely severs the business relationship with the formerly franchised region.

The *Colts* decision illustrates that the likelihood of confusion test still applies to a case of old sport franchise name which is no longer actively used in commerce. In *Colts* the court emphasized that a subsequent user of a bygone trademark is required to take reasonable precautions to prevent consumer confusion (McCarthy, 2011).

According to the decision, a subsequent user may not use an old mark, even if it has been abandoned, in a way that would likely create an impression that the old team is still in operation or the subsequent user is somehow affiliated with the old franchise. In addition to this substantive ruling, *Colts* specifically articulated that the reliability of survey evidence would be dispositive in a trademark infringement claim where the proof of likelihood of confusion is a central issue (Brogan, 1996). Indeed, *Colts* shows how far a likelihood of confusion claim may go under the notion of continuing commercial impression and the importance of valid consumer survey evidence (Mark & Jacoby, 2006).

For purposes of this study, the *Colts* case raised a central issue regarding the use of survey evidence in sports trademark litigation. It is the only sport trademark case to emphasize both the role and permissibility of survey evidence in the court's decision. The sharp criticism of the survey evidence posed by the court in *Colts* further emphasizes the critical role of survey evidence in trademark litigation. The case also highlighted the value of scientifically sound survey evidence to such type of cases. Thus, the *Colts* case is the most desirable test case for examining and illustrating the significance of new statistical techniques to assess the validity of survey evidence related to consumer confusion. The lack of clearly identified criteria for the use of consumer survey evidence and the inconsistent outcomes in sport trademark litigation illustrates the importance of examining the reliability of survey evidence in light of *Colts* and subsequent sport trademark litigation.

Analysis of the Survey Evidence Used in *Colts*

In *Colts*, Baltimore CFL Colts had to deal with a critical problem. The NFL's survey data successfully presented that a substantial number of consumers were confused when they encountered "Baltimore CFL Colts" and mistakenly believed that the CFL team would be somewhat affiliated with or related to the old Colts that had moved to Indianapolis. Nevertheless, Baltimore could contend that the predominant source of the consumer confusion would not be the term "Colts," but rather the term "Baltimore" alone, or alternatively, "Baltimore Colts" as a nontransferable cultural institution (See *MLB Properties, Inc. v. Sed Non Olet Denarius, Ltd.*, 1993). If the term "Baltimore" was indeed a predominant factor for the consumer confusion, it could have seriously weakened the NFL's position in the lawsuit because trademark law generally does not grant exclusive rights associated with a geographical name without a secondary meaning (*Zatarian's, Inc. v. Oak Grove Smokehouse, Inc.*, 1983). Hence, if the consumer confusion was mainly connected to the name "Baltimore," the NFL might have been forced to persuade the court that not only did they exclusively own the mark "Colts," but also their rights associated with "Baltimore Colts" had not extinguished at all even after the team left Baltimore ten years before the litigation and failed to renew the federal registration of the mark.

To counter this potential problem, the NFL's expert incorporated a controlling stimulus into the survey protocols to demonstrate that the source of the confusion would likely be "Colts," not "Baltimore" alone or "Baltimore Colts" as a schematically inseparable term. To eliminate the possibility that the source of the consumer confusion was "Baltimore," the expert implemented "Balti-

more Horses" as a controlling stimulus. In one survey protocol, consumers were exposed to "Baltimore CFL Colts" and "Baltimore/Indianapolis Colts." Then a questionnaire measured whether they were confused by mistakenly believing that "Baltimore CFL Colts" was somewhat related to "Indianapolis Colts" or its bygone predecessor, "Baltimore Colts." In the other protocol, consumers were instead given "Baltimore Horses" and the identical questionnaire measured the degree of consumer confusion. The data analysis following the separate protocols showed that a substantial number of consumers were confused when they were encountered "Baltimore CFL Colts" but no confusion with the pair of "Baltimore Horses" v. "Baltimore/Indianapolis Colts." Based on the protocols, the NFL's expert made an inference that the source of the consumer confusion is not "Baltimore," but "Colts." Not surprisingly, the court opinion expressly criticized that the NFL's protocol might have been flawed or at least somewhat biased because most consumers would hardly believe that the extremely unappetizing name, the "Horses," would be a real trademark or a legitimate candidate for an NFL franchise (*Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Ltd. Partnership*, 1994, p. 415). Nevertheless, the court gave the NFL's evidence a considerable weight seemingly because the defendant had failed to raise any persuasive challenge to refute the biased evidence.

The defendant's expert witness in *Colts* could have challenged the face validity of the measurement procedure (Lipton, 1987) without any statistical analysis or other forms of scientific inquiry. For instance, the expert might contend that "Colts" as a schematic construct would never be completely separable from "Baltimore" in the given measurement protocol because "Baltimore" and "Colts" in the given context would be schematically intertwined and already merged into a single "non-transferable cultural institution" (See, *MLB Properties, Inc. v. Sed Non Olet Denarius, Ltd.*, 1993). In fact, two words had been conjoined together for a substantial period of time in marketing efforts as well as broadcasts so that most consumers would likely perceive two terms as one inseparable schematic entity in relation to professional football. Hence, to identify the predominant source of the consumer confusion at issue, the plaintiffs should have controlled at least most of the informational cues that automatically associate "Baltimore" with "Colts." As a result, the plaintiffs should have employed far more sophisticated controlling protocols such as a double-blind setup with Solomon Four-Group Experimentation (Lipton, 1987; Thomas, Nelson, & Silverman, 2011) rather than the one-dimensional comparison between "Baltimore CFL Colts" and "Baltimore Horses."

In addition to attacking the face validity of the measurement protocols, the defendant's expert could have presented a scientifically grounded argument by conducting a multivariate statistical analysis with respect to the plaintiffs' survey data. While it could be intuitively surmised that "Baltimore Horses" would be a biased controlling stimulus, the scientific proof demonstrating that the stimulus is indeed statistically unacceptable in light of the employed protocols must have been far more persuasive and powerful. Basically, "Horses" and "Colts" in the situation might not be concurrently measured and compared by using the identical questionnaire because they would be completely different schematic properties in the situation dealing with a probable name of a professional sport franchise. Who will choose "Horses" for an NFL team? Statistically, a structural analysis such as the Congenerity Test (Ohanian, 1990) might have revealed that "Horses" in the

given context cannot be compared with "Colts" under the same survey instrument intending to measure consumer confusion because they are schematically different constructs. Even if Colts and Horses are both farm animals and could easily coexist on a ranch, given the context of a probable name for a professional sport franchise, the terms are schematically too different for being a valid comparison. This challenge supported by some coherent and objective scientific data analysis could have helped the Baltimore's position in the lawsuit remarkably.

Commentators and practitioners have pointed out that biased questions in a survey instrument might be one of the most critical flaws in relation to the reliability of survey evidence (Brezina, 2005; Dutka, 1982; Manta, 2007; Redway, 2005; Sarel & Marmorstein, 2002). In particular, a controlling stimulus must not be biased because it is indeed a foundational threshold in the data analysis and manipulation from which the expert will likely approximate the level of consumer confusion (Manta, 2007) or identify some genuine source of consumer confusion as was attempted in *Colts*. Given the critical role of a controlling stimulus in trademark survey protocols, data collectors must present such crucial information in ways similar to the real market conditions where consumers interact with trademarks at issue (Sarel & Marmorstein, 2002). Indeed, the use of "Horses" as the controlling stimulus should have been vigorously challenged. The next section of this paper explores the use of a Congererity Test to screen and/or validate survey evidence in sport trademark litigation.

Congenerity Test as Gatekeeping Methodology in Sport Trademark Litigation

There are diverse statistical methodologies that may assess the reliability of survey evidence introduced in trademark litigation, for instance, Multi-Trait and Multi-Method matrix ("MTMM matrix"), Cronbach's α, and the structural analyses of measurement platforms, so called the *SEM*. Among the enumerated techniques, generally speaking, the MTMM matrix would administer the most parsimonious and vigorous construct validation procedures with respect to a given trademark survey. Unfortunately, its practical utility in litigation might be somewhat limited mainly because the MTMM matrix requires at least two completely different measurement schemes that concurrently capture the same information (e.g., a simultaneous measurement of likelihood of confusion based on a survey as well as some experimental settings). Although Cronbach's α is a product of simple psychometrical equations that easily estimates the reliability of individual measurement items in a survey instrument (i.e., respective questions), the index may not be suitable for the evaluation of a survey instrument in its entirety. Moreover, Cronbach's α does not have any capacity to administer a series of systematic analyses with respect to a given survey protocol, which is necessary for the comparability test with respect to two independent schematic properties such as aforementioned "Horses" and "Colts." Meanwhile, a structural analysis (i.e., the *SEM*) can perform various tests of instrumental validation as well as the structural analyses available for the comparability test. For instance, a structural analysis such as Congenerity Test (Ohanian, 1990) may examine whether the "Horses" implemented as a controlling stimulus in *Colts* is psychometrically identical with "Colts" so that they can be concurrently measured and compared in the given protocols.

Validating Survey Evidence by Using Structural Analysis

The *SEM* is a multivariate technique basically combining the matrices of the Confirmatory Factor Analysis and causal modeling. It is one of the most widely used statistical techniques in modern social science research. In fact, the *SEM* is an effective technique in dealing with two fundamental problems of scientific inferences in social science: (1) the measurement of constructs, and (2) the explanation of the causalities among latent variables (Bagozzi, 1980). Since the main subject of this article is not related to matters of causality but measurement of constructs (e.g., "Horses," vs. "Colts"), it exclusively focuses on the Confirmatory Factor Analysis which is designed to address the issue of measurement error in social science.

Measurement error, so called "disturbances" in psychometrical parlance, is basically endemic in a social science methodology because it mostly deals with an unobservable latent variable that may not be directly measurable (e.g., likelihood of confusion). For instance, when an expert in a trademark infringement case attempts to conduct a survey for the plaintiff, the questionnaire cannot directly measure the degree of the likelihood of consumer confusion at issue. In fact, the investigator indirectly measures the concept (i.e., likelihood of confusion) by asking questions that theoretically approximate the level of the consumer confusion. For example, the survey questionnaire would ask respondents "Do you know what league the team is in?" after respondents are shown two pairs of terms: "Baltimore CFL Colts" v. "Baltimore/Indianapolis Colts" and "Baltimore Horses" v. "Baltimore/Indianapolis Colts." Because this level of likelihood of confusion is indirectly measured, it is an unobservable latent variable. Since the latent variable is represented by its subitem questions only based on some theoretical or conceptual basis, a perfect configuration of the latent variable might be impossible in reality. The Confirmatory Factor Analysis has been devised to deal with this problem. The procedure primarily indicates how effectively a set of questions captures the psychological sphere of the corresponding latent variable. It is one special type of *SEM* that tests the construct validity of a measurement platform. Particularly, it investigates: (1) whether latent variables in a proposed measurement model are properly represented by their subitems (convergent validity); (2) whether multiple latent variables in a platform are distinctive constructs (discriminant validity); and (3) whether a proposed measurement model as a whole is statistically acceptable (model fitness).

In a trademark infringement case, a Confirmatory Factor Analysis can scrutinize the instrumental validity of the survey questionnaire administered by a litigant. It can statistically approximate and evaluate how much weight the fact finder would give to the survey evidence presenting the proof of consumer confusion.

Congenerity Test: Are Two Schematic Constructs the Same Psychological Properties?

Congenerity Test is a statistical technique employed by social scientists in investigating whether two different psychological properties can be concurrently measured and compared by using an identical measurement scheme (e.g., a survey questionnaire). In communication science, the method was used to examine whether a measurement scale can effectively capture schematic images of two different

spokespersons in celebrity endorsement settings (Ohanian, 1990). On the other hand, a sport marketing study has implemented such technique to test whether a pairing of sport event with a sponsor in a sponsorship program can be psychometrically matched by using a standardized survey instrument (Cho & Kang, 2012). Cho and Kang (2012) tested the comparability of Aaker's brand personality scale (Aaker, 1997) with respect to 2×2 pairs of sport events (Professional Golfers' Association and Major League Baseball) and major sponsors (Mercedes Benz and Nike). The tested measurement scale consists of five dimensions (sincerity, excitement, competence, sophistication, and ruggedness). The study indicated that the chosen sport events and sponsors may not be psychometrically comparable in terms of at least one dimension of the scale for all four pairs. That is, while various sponsorship studies attempt to compare sport events and sponsors in terms of the Aaker's five dimensions, some matching efforts must have been questionable psychometrically.

If a Congenerity Test shows that two schematic entities are psychometrically comparable, it establishes an inference that they possess the same psychological property (i.e., same schemata in the given context). In contrast, if the Congenerity Test result is negative, the two entities are different properties and may not be concurrently measured or psychometrically compared in a given protocol. In *Colts*, the defendant could have used multivariate statistics to demonstrate that incorporating "Horses" in the survey protocols as a controlling stimulus might be significantly biased. In other words, the "Horses" and "Colts" are different schematic entities probably because few people with reasonable mind would perceive that the "Horses" may actually be or is a conceivable option for a real professional sport franchise.

In fact, there are two seemingly related but different types of inquiries that would be examined by the structural analysis of a Congenerity Test. First, Congenerity Test may investigate the external validity of a measurement platform. External validity refers to the generalizability of the survey instrument. In this regard, the test examines whether different versions of the measurement scale can equally capture the sphere of one psychological configuration. In a trademark infringement claim, a plaintiff may have to use two different versions of a survey instrument for measuring consumer confusion (e.g., English and Spanish) if a substantial number of typical buyers in the relevant market are Hispanics. Then the generalizability of the measurement scale would become an issue. In this hypothetical situation, a Congenerity Test may examine whether both an English and Spanish version of the scale accurately measure the consumer confusion after surveying a bilingual population within the relevant market.

Secondly, being mentioned, the Congenerity Test can examine the comparability of two different constructs (i.e., whether two entities share the same type of psychological properties). The test assesses the comparability of different constructs by estimating the χ^2 difference between a constrained model and a base model of a given measurement platform. A model χ^2 indicates the extent to which a proposed measurement model captures targeted schematic domains. Therefore, a Congenerity Test is based on the notion that if χ^2 values obtained from two independent observations for an identical measurement scheme are statistically different, two separate observations likely captured two different psychological properties. In the constrained model for the Congenerity Test, a pair of measurement items is fixed as a single parameter, whereas the same parameters are freed in the base model. If a χ^2 difference between these two models is significant, it is inferred that the constrained

parameters would be measuring psychometrically different constructs. Inversely, if a calculated χ^2 difference is not significant, it statistically warrants that the fixed parameters are measuring an identical psychological property. The procedures of Congenerity Test can be described as (Cho & Kang, 2012):

M_1 (Base Model): λ_{11}. . . . λ_{jk}, δ_1. . . . . δ_k, and ρ_1 & ρ_2 not constrained.
M_2 (Constrained Model): $\rho_1 = \rho_2 = 1$.

While the factor loadings (λ_{11}. . . . λ_{jk}), error variances (δ_1. . . . . δ_k), and factor variances of two constructs (ρ_1 & ρ_2) in the M_1 (Base Model) are not fixed as same, such parameters in M_2 are all constrained as same. If the χ^2 difference between M_1 and M_2 is not significant, the imposed parameter constraint (i.e., $\rho_1 = \rho_2$) is then statistically acceptable. It implies that the parameter ρ_1 and ρ_2 are connected to the same psychological property and they are concurrently comparable in a single context, as a result. In contrast, if the χ^2 difference is significant, the imposed constraint of the M_2 is statistically rejected. The result implies that the said parameters (i.e., ρ_1 and ρ_2) cannot be concurrently measured and psychometrically compared because they are different schematic properties.

In *Colts*, the structural analysis of a Congenerity Test could have revealed that "Baltimore Horses" may not be concurrently measured and compared with "Baltimore/Indianapolis Colts" in the given context by using a series of identical questions. That is, few people would have perceived the "Horses" similar to the "Colts" in the administered protocols. As such, the technique might be a parsimonious gatekeeping technique in trademark litigation where a litigant introduces a seemingly biased controlling stimulus and the comparability of different words is a dispositive issue. One of the major shortcomings of Baltimore's challenge in *Colts* was the lack of scientific examination of the NFL's survey data that could have revealed some critical flaw in the implemented measurement protocols. Particularly, a scientifically grounded argument challenging the use of the "Horses" as a controlling stimulus could have been effective. It might have seriously weakened the credibility of the NFL's survey evidence. Nevertheless, the court opinion does not mention any scientifically grounded challenge raised by Baltimore's expert regarding the survey evidence. Instead, according to the opinion, the defendant's expert fruitlessly pointed out that survey methodology mainly belongs to sociology rather than psychology, which might not be persuasive at all, specifically to Justice Posner who is presumably very conversant with the turf war among different social science disciplines. Given the absence of scientifically supported argument, a Congenerity Test could have provided the defendant with a far more coherent argument attacking the NFL's potentially biased survey information. In sum, the structural analysis of a Congenerity Test could have demonstrated that the plaintiffs' survey data would be unreliable. It might have demonstrated that "Horses" must not be used as a controlling stimulus to preclude the possibility that the consumer confusion manifested in the survey data would likely come from "Baltimore" or at least something closely intertwined with "Baltimore" instead of "Colts."

Hypothetical Application of a Congenerity Test in *Colts*

The best scenario for this study is certainly using the actual survey evidence introduced by the parties in the *Colts* case. Unfortunately, the data are not publicly available

mainly due to the confidentiality agreements among parties and their experts. Although a hypothetical data analysis would not be the best option, it could be still meaningful in consideration of the fact that the main purpose of this article is the introduction of the technique rather than a substantive analysis of the real data. Hence, this article now demonstrates a hypothetical situation where Baltimore's expert in *Colts* raises a statistically supported objection to the survey data introduced by the NFL. Specifically, the demonstration presents how the technique could have supported Baltimore in challenging the use of the "Horses" as a controlling stimulus in the survey protocols to infer that the source of the consumer confusion at issue could be the term "Baltimore" instead of "Colts."

According to the case, the NFL's survey protocol initially presents respondents with various merchandise items bearing two sets of trademarks [i.e., "Baltimore CFL Colts" v. "Baltimore/Indianapolis Colts" ("Colts pair") and "Baltimore Horses" v. "Baltimore/Indianapolis Colts" ("Horses pair")]. Next, respondents are given a series of questions measuring the likelihood of confusion allegedly attributed to the defendant's mark, "Baltimore CFL Colts" or the controlling stimulus, "Baltimore Horses." The questions in this hypothetical situation are presented in Table 1.

The NFL's survey data now indicate that a substantial number of respondents are not confused with respect to the "Horses pair" (i.e., "Baltimore Horses" v. "Baltimore/Indianapolis Colts"), while being confused by the "Colts pair" (i.e., "Baltimore CFL Colts" v. "Baltimore/Indianapolis Colts"). In response to this

### Table 1   Hypothetical Survey Questions

Measurement items (Qs)	Latent variables
(Q #1-a) "Do you know what sport this team plays?"	Awareness of mark
(Q #1-b) "If yes, what sport?"	Awareness of mark
(Q #1-c) "Do you know any player's name in this team?	Awareness of mark
(Q #2-a) "Do you know what teams it plays against?"	Awareness of competitors
(Q #2-b) "If yes, list opponents you know up to three"	Awareness of competitors
(Q #2-c) "Do you know any opponent team's player?"	Awareness of competitors
(Q #3-a) "Do you know what league the team is in?"	Awareness of league
(Q #3-b) "If yes, what league?"	Awareness of league
(Q #4-a) "Does the team need someone's permission to use the name?"	TM right for team's name
(Q #4-b) "If yes, who is it?"	TM right for team's name
(Q #5-a) "Does the team need someone's permission to use the symbol?"	TM right for team's symbol
(Q #5-b) "If yes, who is it?"	TM right for team's symbol

potentially biased evidence of consumer confusion, Baltimore's expert now conducts the Congenerity Test to demonstrate that the "Horses" and "Colts" in the survey protocols might not likely be the same psychological property because most respondents may perceive "Horses" and "Colts" very differently in relation to the administered context. Therefore, Baltimore's expert seeks to indicate that "Horses" and "Colts" may not be concurrently measured and psychometrically compared in terms of the same questionnaire. Here the expert ultimately tries to refute the reliability of the survey evidence arguing that the source of the consumer confusion is not "Colts" but instead is the term "Baltimore." Of course, if the outcome of the test indicates that two constructs are statistically comparable, the only ground for the defendant's challenge would be the previously mentioned nonstatistical argument. In fact, all survey evidence and data analyses challenging opponents' survey evidence in trademark litigation would be a double-edged sword because they are always discoverable once conducted. It is a strategic or tactical decision to be made by parties and experts in the proceedings. Nevertheless, given the parsimonious nature of the Congenerity Test, it must have been a worthwhile tactical move for the defendant in *Colts*.

Assuming that the defendant in the hypothetical situation has decided to conduct the Congenerity Test, following statistical procedures will be implemented. To implement the test, two separate measurement platforms capturing "Horses" and "Colts" must be juxtaposed and consolidated as one measurement solution. Therefore, two identical sets of the questions measuring consumer confusion with respect to "Horses" and "Colts" are merged into one measurement solution (Figure 1). This consolidated solution is the base model for the Congenerity Test. All diagrams presented in this article are hypothetical. They are presented only for the purpose of the visual demonstration of the testing procedures rather than any substantive data analysis.

The next step is running a Confirmatory Factor Analysis for the base model (Figure 2) in which no parameters are constrained ($\lambda_{11}. \ldots \lambda_{jk}, \delta_1. \ldots \ldots \delta_k$, and ρ_1 & ρ_2 not constrained). It will estimate a model χ^2. The χ^2 differences between this base model and five constrained models where the corresponding sets of the questions are fixed as same parameters ($\rho_1 = \rho_2 = 1$) ultimately indicate whether "Horses" and "Colts" are the same psychological property in the given context. In sum, the base model (Figure 2) does not statistically assume that "Baltimore CFL Colts" and "Baltimore Horses" are psychometrically identical while constrained models assume that two are the same with respect to a particular dimension of the administered solution. The model χ^2 of the base model is compared with constrained models sequentially.

To construct the first constrained model, the first set of corresponding parameters (i.e., Q #1-a, b and c) with respect to "Baltimore CFL Colts" and "Baltimore Horses" are fixed as the same parameters (Figure 3). This constrained model assumes that two entities are the same schematic constructs with respect to the questions. If the χ^2 difference between this constrained model and the base model is statistically significant, the hypothesized constraint, $\rho_1 = \rho_2 = 1$, must be rejected with respect to the three questions (i.e., Q #1-a, b and c), meaning that "Colts" and "Horses" in the administered measurement protocols may not be considered the same psychological property in terms of the first latent variable, "awareness of mark, represented by the three questions" (See Table 1). The results indicate that
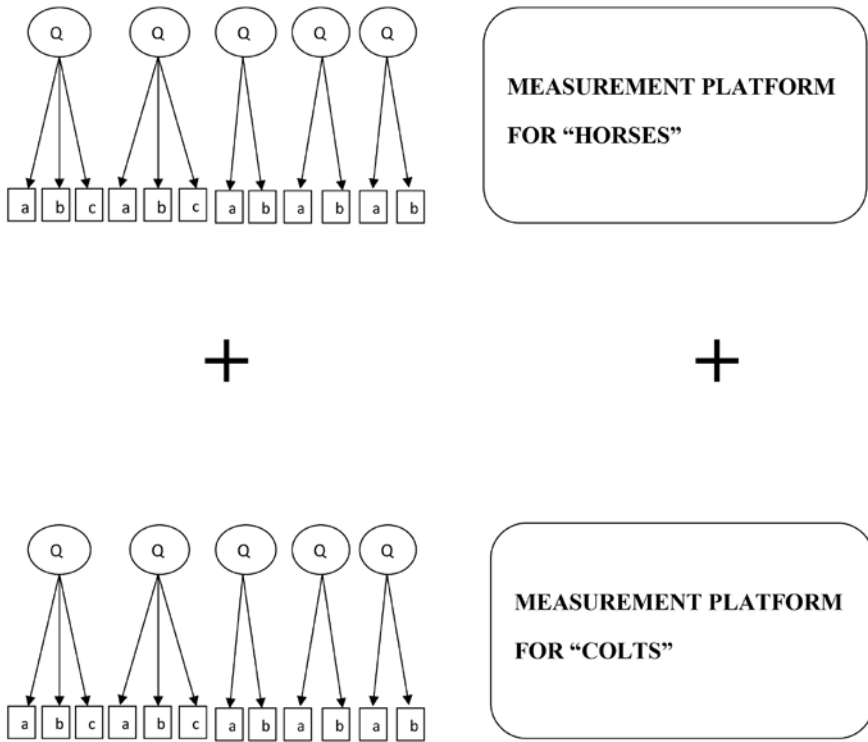
**Figure 1** — Consolidation of measurement platforms.

"Horses" might be a biased controlling stimulus, which cannot be concurrently measured and compared with "Colts" in terms of "awareness of mark" represented by Q #1-a, b and c.

   The same procedures are repeated for the remaining sets of questions to examine whether "Horses" can be concurrently measured and compared with "Colts" in terms of the other four latent variables (i.e., "awareness of competitors," "awareness of league," "TM right for team's name," and "TM right for team's symbol") (Table 1). If a majority of the constrained models indicates that the two schematic entities are not the same, the Congenerity Test would likely refute the NFL's survey evidence, in particular, the use of the "Horses" as a controlling stimulus in the protocols.

Conclusion

As contemporary mark owners seek more aggressive protection for their trademark rights (McKelvey, Fairley, & Groza, 2010), trademark infringement claims will remain as a primary legal recourse for mark owners and proof of likelihood of confusion will remain essential to the success of those claims. According to a seminal U.S. Supreme Court decision, *McLean v. Fleming* (1878), trademark
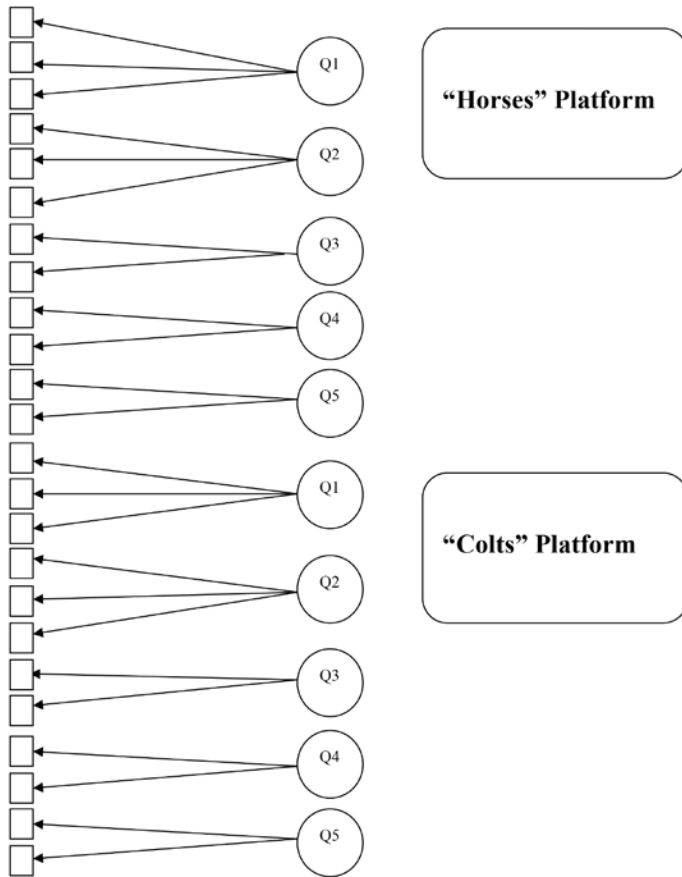
Figure 2 — Congenerity test: base model (no parameters are constrained).

infringement actions are consumer-oriented legal claims (Caughey, 1956). The opinion declared that the likelihood of confusion test must focus on the examination of ordinary consumers' deception. Under *McLean*, it is presumed that the likelihood of confusion test may not be entirely dependent upon an expert's opinion. The legal test is whether a substantial number of consumers in the relevant market would likely be confused as to the sources or affiliation of commodity or service because of the allegedly infringing mark. Thus, survey evidence will continue to play a critical role in trademark litigation as discussed in this paper. In spite of this remarkable significance, Thornburg (2005) found that some courts are still hesitant in allowing survey evidence even when the factual question of consumer confusion is obviously a central issue. One study (Beebe, 2006) also indicates that survey evidence is not a statistically significant determinant indecisions of trademark infringement claims where cases are examined based on the multifactor likelihood of confusion analysis. It epitomizes the ambivalent view of courts toward
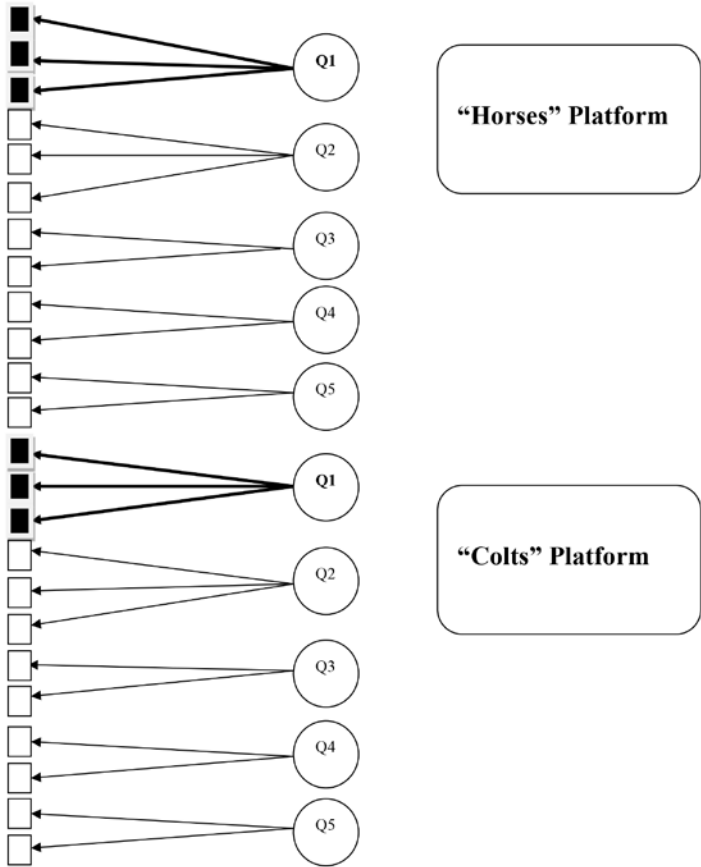
Figure 3 — Constrained model (Q #1-a, b, c are fixed as same parameters).

the legal utility of survey evidence. While it is true that a consumer survey may deliver crucial information in trademark infringement cases, there is no consensus as to a gatekeeping methodology that may objectively estimate the reliability of survey information. This article demonstrates that a Congenerity Test might be a possible gatekeeping technique implemented in trademark litigation specifically where the psychometrical comparability of two different schematic constructs is a material issue. Although a structural analysis such as Congenerity Test may not be a magic wand to screen out all unreliable survey information, the technique at least provides some guidelines and starting points in pursuit of more reliable measurement schemes in survey procedures that would reduce biased data like the data introduced in *Colts*.

This article suggests several future studies. First of all, while this paper discussed the structural analysis in relation to the likelihood of confusion claim only, the same or similar technique might provide some valuable information in cases of

secondary meaning disputes as well. Secondary meaning is another critical aspect in the jurisprudence of trademark law and consumer surveys are frequently employed in the inquiry. Therefore, the Congenerity Test or other structural technique might provide gatekeeping protocols to screen out some unreliable information as demonstrated here. Secondly, a study may investigate the relationship between the multifactor likelihood of confusion test and fair use defense. In *KP Permanent Make-up v. Lasting Impressions* (2004), the U.S. Supreme Court decided that a defendant in a trademark infringement action may resort to fair use defense even if there is some evidence of likelihood of confusion established by the plaintiff. Given the significant similarities between the fair use defense and the multifactor likelihood of confusion analysis (Greene, 2006), an empirical study may examine how the fair use defense in trademark infringement is correlated with the test of consumer confusion by using an outcome determinant analysis. Thirdly, a series of consumer surveys would be conducted to seek more coherent empirical substances in the area of trademark dilution. While the cause of action for trademark dilution is provided under the federal statute, the claim's underlying public policy has been controversial because it is somewhat departure from the traditional notion of consumer-oriented trademark law. Most of all, it has not been empirically examined whether blurring or tarnishment of famous trademarks would actually be detrimental to distinctive values of famous marks. Given the dearth of relevant literature, an empirical study employing *SEM* technique would test a causal relationship between blurring or tarnishment activities in the market and presumed negative effects on the distinctive brand value of famous marks.

References

15. U.S.C. § 1114(1)(a) (West 2005).

15. U.S.C. § 1125(a)(1) (West 2005).

Aaker, D.A. (1996). *Building Strong Brands*. New York, NY: The Free Press.

Aaker, J.L. (1997). Dimensions of brand personality. *JMR, Journal of Marketing Research, 34*(3), 347–356. doi:10.2307/3151897

Angel, A. (2010). Let go my logo. *ESPN Magazine, 13*(8), 40.

Babington, C., & Denlinger, K. (1995, November 7). Modell announces Browns' move to Baltimore. The Washington Post. Retrieved October 12, 2011 at http://www.washingtonpost.com/wp-srv/sports/longterm/memories/1995/95nfl4.htm

Bagozzi, R.P. (1980). *Causal Models in Marketing*. New York, NY: John Wiley and Sons, Inc.

Beebe, B. (2006). An empirical study of the multifactor tests for trademark infringement. *California Law Review, 94,* 1581–1654. doi:10.2307/20439078

Brezina, D.C. (2005). Survey types in likelihood-of-confusion cases: Examining confusion in competitive and noncompetitive products. *Intellectual Property& Technology Law Journal, 17*(9), 1–7.

Brogan, S. H. (1996). Who are these "Colts?": The likelihood of confusion, consumer survey evidence and trademark abandonment in Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club, Ltd. *Marquette Sports Law Journal, 7*, 39–75.

Caughey, R.E. (1956). The use of public polls, surveys and sampling as evidence in litigation, and particularly trademark and unfair competition cases. *California Law Review, 44,* 539–546. doi:10.2307/3478753

Cho, S., & Kang, J. (2012). Psychometric comparability of Brand Personality Scale: Assessing brand personality matching between sports and corporate sponsors by using Congenerity Test. *International Journal of Sports Marketing & Sponsorship, 13*(4), 295–312.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 573 (1993).

Dutka, S. (1982). Law office memos: The use of survey research in legal proceedings. *American Bar Association Journal. American Bar Association, 68,* 1508–1510.

Eagle Snacks, Inc. v. Nabisco Brands, Inc., 625 F.Supp. 571 (D.N.J. 1985).

Geise, S.N. (1995). A whole new ballgame: The application of trademark law to sports mark litigation. *Seton Hall Journal of Sport Law, 5,* 553–574.

Grady, J., & McKelvey, S. (2008). Trademark protection of school colors: *Smack Apparel* and *Sinks* decisions trigger *color*-ful legal debate for the collegiate licensing industry. *Journal of Legal Aspects of Sport, 18*(2), 207–242.

Greene, S. (2006). Sorting out "fair use" and "likelihood of confusion" in trademark law. *American Business Law Journal, 43*(1), 43–77. doi:10.1111/j.1744-1714.2006.00011.x

Heath, T. (2006, February 22). Court battle with apparel firm threatens Nats merchandise. The Washington Post. Available at: http://www.washingtonpost.com/wp-dyn/content/article/2006/02/21/AR2006022101721.html

Hylton, J.G. (2011). The over-protection of intellectual property rights in sports in the United States and elsewhere. *Journal of Legal Aspects of Sport, 21*(1), 43–73.

Indianapolis Colts, Inc. v. Mayor and City Council of Baltimore, 775 F.2d 177 (7th Cir. 1985).

Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Limited Partnership, 34 F.3d 410 (7th Cir. 1994).

Jones, L.C. (1989). Developing and using survey evidence in trademark litigation. *Memphis State University Law Review, 19,* 471–490.

Kahn, M.A. (2004). May the best merchandise win: The law of non-trademark uses of sports logos. *Marquette Sports Law Review, 14,* 283–317.

Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).

KP Permanent Make-up v. Lasting Impressions, 543 U.S. 111 (2004).

Lemley, M.A. (1999). The modern Lanham Act and the death of common sense. *The Yale Law Journal, 108,* 1687–1715. doi:10.2307/797447

Leone, K.C. (1997). No team, no peace: Franchise free agency in the National Football League. *Columbia Law Review, 97,* 473–523. doi:10.2307/1123368

Lipton, J.P. (1987). Trademark litigation: A new look at the use of social science evidence. *Arizona Law Review, 29,* 639–665.

Major League Baseball Properties, Inc. et al. v. Bygone Sports. LLC, (2005, June 13). Complaint filed in the Southern District of New York (CV-05526).

Manta, I.D. (2007). In search of validity: A new model for the content and procedural treatment of trademark infringement surveys. *Cardozo Arts &Entertainment Law Journal, 24,* 1027–1072.

Mark, G., & Jacoby, J. (2006). Continuing commercial impression: Applications and measurement. *Marquette Intellectual Property Law Review, 10,* 433–458.

Medici Classics Prods, LLC. v. Medici Group, LLC., 590 F.Supp.2d 548 (S.D.N.Y. 2008).

McCarthy, J.T. (2011). *McCarthy on trademarks and unfair competition* (4th ed.).

McKelvey, S., Fairley, S., & Groza, M. (2010). Caught in the web?: The communication of trademark rights and licensing policy on university official athletic websites. *Journal of Legal Aspects of Sport, 20*(1), 1–33.

McKenna, M.P. (2007). The normative foundations of trademark law. *The Notre Dame Law Review, 82*(5), 1839–1916.

McLean v. Fleming, 96 U.S. 245 (1878).

Medici Classics Prods, LLC. v. Medici Group, LLC., 590 F. Supp. 2d 548 (S.D.N.Y. 2008)

MLB Properties, Inc. v. Sed Non Olet Denarius, Ltd., 817 F. Supp. 1101 (S.D.N.Y. 1993).

Nationals settle trademark lawsuit. (2006, March 14). *Sports Illustrated Online* (SI.com). Available at: http://sportsillustrated.cnn.com/2006/baseball/mlb/wires/03/14/2010.ap.bbn.nationals.trademark.0220/

Notes of Advisory Committee, Fed. R. Evid. 702 (West 2005).

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39–52. doi:10.1080/00913367.1990.10673191

Playboy Enterprises, Inc. v. Netscape Communications Corp., 55 F. Supp. 2d 1070(C.D. Cal. 1999)

Polaroid Corp. v. Polarad Electronics, Corp., 287 F.2d 492 (2d Cir. 1961).

Redway, H.J. (2005). The use of survey evidence in trademark and unfair competition cases: Insights and practice tip. *The Computer and Internet Lawyer, 22*(10), 11–15.

Sarel, D., & Marmorstein, H. (2002). Designing confusion surveys for cyberspace trademark litigation: The admissibility v. weight debate. *Intellectual Property& Technology Law Journal, 14*(9), 12–17.

Thomas, J.R., Nelson, J.K., & Silverman, S.J. (2011). *Research methods in physical activity*. Champaign, IL: Human Kinetics.

Thornburg, R.H. (2005). Trademark survey evidence: Review of current trends in the Ninth Circuit. *Santa Clara Computer and High-Technology Law Journal, 21,* 715–743.

Wheble, K.M. (2009). Evaluating anecdotal evidence of confusion. *Intellectual Property Litigation, 21*(1), 12–14.

Zatarian's, Inc. v. Oak Grove Smokehouse, Inc., 698 F.2d 786 (5th Cir. 1983)

Zippo Manufacturing Co. v. Rogers Imports, Inc., 216 F. Supp. 670 (1963).

Appendix: Glossary of Statistical Terms Used in the Article

Cronbach's α: A statistical index estimating how effectively a measurement item captures a sphere of psychometrical property.

Congenerity Test: A type of Structural Equation Modeling that primarily analyzes comparability of two schematic constructs in a single measurement platform.

Manipulation check: A scientific procedure to checks whether an observed result is a product of implemented treatments. It is mostly used in an experimental setting.

Multi-Trait and Multi-Method matrix ("MTMM" matrix): A methodology that checks the reliability of a measurement scheme based on multiple types of observation.

Parameters: Unknown variances estimated by data entered by a researcher. In a Congenerity Test, the lambda (λ) coefficient indicates a factor loading from a latent variable to an observed item while the theta-delta (δ) stands for an error variance of an observed item. The phi (ρ) designates a variance of a latent variable.

Psychometrics: A methodological notion that quantifies psychological valences by assigning numeric values to psychological properties. To assign quantified values to various psychological manifestations (e.g., attitude, schemata, intention, etc.), researchers frequently use a scale such as Likert scale.

Structural Equation Modeling (*SEM*): A statistical procedure based on factor analysis and causal modeling designed to investigate relationships among latent variables.