

Surveys on a Tightrope — the Convergent Validity Net

Ruth M. Corbin & Fiona Isaacson*

Survey evidence has been used for more than half a century in intellectual property litigation in Canada. Frequency of acceptance of such evidence has dipped below 20 per cent in the last five years. The 2011 Supreme Court decision in Masterpiece has ignited a debate about when, and under what circumstances, surveys will continue to be welcomed by courts. Social science and modern technology offer under-used alternatives to surveys that can be combined in strong ways. Results of alternative research approaches produce an integrated snapshot of the marketplace, underpinned by the principle of “convergent validity.” This principle is consistent with the style of new knowledge development in several other fields, and offers broad-based assurance of reliability and validity of market evidence.

Les sondages sont utilisés en preuve depuis plus d'un demi-siècle dans les litiges en matière de propriété intellectuelle au Canada. La fréquence d'admissibilité de ce genre de preuve est tombée en deçà de 20 pour cent au cours des cinq dernières années. L'arrêt Masterpiece de la Cour suprême en 2011 a provoqué un débat à savoir quand et dans quelles circonstances les sondages peuvent être utilisés devant les tribunaux. Les sciences humaines et la technologie moderne offrent des solutions de rechange trop peu utilisées, par rapport aux sondages, qui peuvent être combinés de façon intéressante. Les résultats obtenus par ces autres méthodes de recherche dressent un portrait intégré du marché et se fondent sur des principes de « validité convergente ». Ce dernier principe est conforme au style de développement de nouvelles connaissances dans plusieurs autres domaines et offre d'importantes garanties de fiabilité et de validité en tant que preuve du marché.

1. INTRODUCTION

Surveys have been used to deliver evidence in intellectual property litigation for more than half a century. Despite the fact that surveys are inherently “hearsay,” being reports of what other people have said, they came to be recognized as permissible exceptions to the hearsay rule when presented by qualified experts.¹ The

* Dr. Ruth M. Corbin is Managing Partner, CorbinPartners Inc., and Adjunct Professor, Osgoode Hall Law School. Fiona Isaacson is Project Manager, CorbinPartners Inc.

¹ See R.M. Corbin & A.K. Gill, *Survey Evidence and the Law Worldwide* (LexisNexis, 2008) ch. 1, for a description of turning point cases in various countries for acceptance of survey evidence.

Schenley case² has generally been recognized as the turning point in Canadian judicial attitude, whereby a properly conducted survey was treated as readily helpful to a determination of the issues. A definitive endorsement of survey evidence was colourfully given in the Canadian Federal Court by Justice Macfarland, writing in *Sun Life*:

Without such evidence, how am I to otherwise determine whether there is likely to be confusion . . . what I think personally is immaterial . . . To attempt to make such a determination without regard to evidence of what others may think or have said would to my mind be nothing more than an exercise in pure judicial fantasy and of not much assistance at all.³

The frequency of survey evidence in case law since the time of *Sun Life* illustrates that the benefits of surveys, when conducted to a high standard, have been considered by courts to outweigh their limitations.

The continued confidence in survey evidence is being tested. Since 2008, surveys have been accepted or relied upon on fewer than one in five occasions when tendered to Canadian courts;⁴ courts and the Trade-marks Opposition Board have sometimes made a point of saying that a survey would be unnecessary in resolving a point of dispute.⁵ In its 2011 *Masterpiece*⁶ trade-marks decision, the Supreme Court of Canada encouraged parties to consult at the case management stage to decide whether survey evidence would be truly useful to the court, as opposed to providing merely an unnecessary dose of “common sense.” Experts, said the Supreme Court, should only be permitted to testify if their testimony is likely to be outside the experience and knowledge of the judge. The view that a single judge might reasonably anticipate the opinions of the broad consumer population appears to stand in stark contrast to Justice Macfarland’s characterization of such a view as “pure judicial fantasy.”⁷ Commenting on the *Masterpiece* case, the Marketing Research and Intelligence Association of Canada (“MRIA”), the national self-regulatory body for survey professionals, advised that “direct interviews with consumers or observations of consumer behaviour remain the most statistically reliable source of conclusions about what the relevant consumer population is likely to perceive when it comes to brands and trade-marks.”⁸ The Supreme Court decision has ignited a debate in the intellectual property community about when, and under what

² *Canadian Schenley Distilleries Ltd. v. Canada’s Manitoba Distillery Ltd.* (1975), 25 C.P.R. (2d) 1 (F.C.T.D.).

³ *Sun Life Assurance Co. of Canada v. Sunlife Juice Ltd.*, 1988 CarswellOnt 926, 20 C.I.P.R. 87, 65 O.R. (2d) 496, 22 C.P.R. (3d) 244 at paras. 20, 21 (H.C.).

⁴ See “Year in Review” compilations of survey evidence in Canada for 2008, 2009, 2010, 2011 at <<http://www.corbinpartners.com/index.php/media-and-publications>>.

⁵ E.g., *Shell Canada Ltd. v. P.T. Sari Incofood Corp.*, 2008 FCA 279; *College of Traditional Chinese Medicine Practitioners & Acupuncturists (British Columbia) v. Council of Natural Medicine College (Canada)*, 2009 FC 1110.

⁶ *Masterpiece Inc. v. Alavida Lifestyles Inc.*, 2011 SCC 27, [2011] 2 S.C.R. 387.

⁷ *Supra* at note 3.

⁸ MRIA, “Member Advisory on Supreme Court of Canada comments on use of Survey Evidence in Brand or Trademark Disputes” posted on its website at <<http://www.mria-arim.ca/NEWS/AdvisorySupremeCourt.asp>>, accessed October 3, 2012.

circumstances, surveys will be welcomed by the courts.⁹ Wherever the dust settles on that debate, the Supreme Court in *Masterpiece* was clear that its standard for survey evidence is “necessity,” and not merely “helpfulness.”

The United Kingdom’s Intellectual Property Office has expressed similar cautiousness with respect to surveys. Its 2012 practice notice¹⁰ announced that survey evidence could be adduced into trade-mark proceedings only with the permission of the hearing officer. The practice notice sets out the criteria which must be satisfied for such evidence to be admitted. Although surveys remain a routine source of evidence in U.S. courts,¹¹ the value of surveys has also recently been challenged in specific U.S. contexts, including disputes concerning trade-mark dilution.¹²

Canada’s Supreme Court was careful to preserve a future for survey evidence for purposes of demonstrating “consumer reactions in the marketplace — exactly the question that the trial judge is addressing in a confusion case”¹³ as long as a proffered survey meets the social scientific standards of reliability and validity. “Reliability” refers to the ability of survey results to be generalized from a sample of consumers to the overall pertinent population. “Validity” refers to whether the right things are measured in the right way.

A rocky road lies ahead for sustaining reliability in surveys. In most situations, perfect statistical reliability is no longer a feasible ideal.¹⁴ Perfect statistical reliability requires a pure random sample of the relevant population. Obtaining a pure random sample from a general consumer audience is impeded by limitations of any of the various methods for administering surveys. For example, door-to-door interviewers cannot plausibly expect to capture a random sample of households. One reason is that apartment buildings are generally inaccessible to uninvited interviewers; certain neighbourhoods that pose safety risks are also *de facto* inaccessible. Mall interviewing captures principally trade-area shoppers in selected cities. Telephone interviewing — once a gold standard for random sampling — is now impeded by blocked-calling controls, answering machines, and households without

⁹ P. Tackaberry, “*Masterpiece v. Alavida*: Supreme Court of Canada Clarifies Likelihood of Confusion Test and Role of Expert Witnesses,” *Trademark Reporter*, 101(5) (2011), 1546–1579.

¹⁰ Tribunal Practice Notice (2/2012) of the Intellectual Property Office, posted online at <<http://www.ipo.gov.uk/pro-types/pro-tm/t-law/t-tpn/t-tpn-2012/t-tpn-22012.htm>>.

¹¹ G.L. Ford, “Intellectual Property Surveys. Annual Cumulative Update 2010,” <<http://www.inta.org/TrademarkBasics/FactSheets/Documents/INTALanhamActSurveysCumulUpdate2010.pdf>>.

¹² R.C. Bird, “The Impact of the *Moseley* decision on Trademark Dilution Law”, 2006, published on the Social Sciences Research Network at <<http://ssrn.com/abstract=903003>>, accessed October 3, 2012.

¹³ *Masterpiece Inc. v. Alavida Lifestyles Inc.*, 2011 SCC 27, [2011] 2 S.C.R. 387 at para. 93.

¹⁴ Advertising Standards Canada acknowledged this reality of sample surveys in its most recently published research guidelines, noting that “perfect random sampling is almost never possible. Accepted industry practices are a permissible substitute.” See “Guidelines for the Use of Comparative Advertising” at <<http://www.adstandards.com/en/ASCLibrary/guidelinesCompAdvertising-en.pdf>>, accessed October 3, 2012.

landlines. Internet surveys, increasingly the cost-effective choice for business applications, are largely served by non-random volunteer panelists.

Yet, despite the infeasibility of true random sampling, survey evidence is likely to withstand challenges to its statistical reliability, for several reasons. The first is that surveys are still heavily relied upon by business and government for significant decisions. That fact has always been a reason why courts have considered them worthy of consideration as inputs to judicial decision-making.¹⁵ Second, the survey industry is evolving to strengthen the quality controls over sample representativeness, respondent identification and interview validation, particularly with respect to Internet surveys. Indeed, the International Organization for Standardization is now making available a recognized "ISO" standard for the market research industry, which will inevitably escalate "best practices."¹⁶ A third reason is that whatever surveys provide by way of enlightenment can be bolstered, where warranted, through integrating other forms of social science evidence, not yet widely exploited by legal professionals. The result of an integrated-methods approach is a portfolio of evidence that is stronger and more convincing than a survey ever would have been on its own. That is the subject matter of the remainder of this article.

2. SURVEYS WERE NEVER THE "EVIDENCE"

Surveys are no more than a methodology for delivering evidence. Despite the possibly-misleading phrase "survey evidence," they were never the evidence itself. Surveys are a delivery vehicle for information about what the public experiences, thinks, believes, or intends to do. They are not the only vehicle. Indeed, in some countries, marketplace evidence has been delivered with little or no reliance on survey methodology.¹⁷

To analyze from first principles what sort of evidence is desirable or necessary, it is worth recalling that "intellectual property" consists of products of the mind. Trade-mark law that protects intellectual property focuses on the consumer's psychological states of mind and thought processes. In the Canadian *Trade-marks Act*, confusion is defined in terms of an "inference," and passing-off is determined by whether the public is "misled."¹⁸ Similar statutory references to consumer belief or perception exist in the trade-mark laws of several countries. Estimating loss of

¹⁵ For example, procedural rule 803(17) in the U.S. Federal Rules of Evidence allows admission of data "generally relied upon by the public or by persons in particular occupations." <http://www.law.cornell.edu/rules/fre/rule_803>, accessed October 3, 2012.

¹⁶ See the official website of the ISO at <<http://www.iso.org/iso/pressrelease.htm?refid=Ref1005>>, accessed October 3, 2012.

¹⁷ INTA, *Types of Evidence Used to Establish Likelihood of Confusion World Survey*. <<http://www.inta.org/Advocacy/Documents/INTATypesofEvidence2009.pdf>>, accessed January 3, 2012. In 2009, the International Trademark Association did a world survey of types of evidence used to establish the likelihood of confusion. Of the 45 countries surveyed, plus the European Court of Justice, just over half were found to accept surveys as evidentiary support. In at least eight countries, the everyday experience or personal opinion of the judge or examiner entered the analysis in a significant way.

¹⁸ At sections 6 and 7, respectively.

future sales arising from infringement or damaged reputation entails a prediction of consumer behaviour. Surveys are one possible source, but not the only source of measuring consumer perception and behaviour anticipated in the language of trademark statutes. Litigants and courts would benefit from any trustworthy social science vehicle that brings in essential evidence applicable to the relevant law.

Accepting surveys as conveyers of relevant evidence means accepting the assumptions on which they are based. These assumptions include, but are not limited to, the assumptions that a random sample of invitations generates a random sample of consumers,¹⁹ that people are willing to tell the truth in surveys, that people are capable of reporting on their own thoughts and intentions, or that the many unavoidable influences on survey answers on any given occasion can be treated as random noise. Surveys, like any method of social science, provide enlightenment only as far as their underlying assumptions will allow.

The same is true for other methods. They may provide enlightenment from a different approach, based on different assumptions. Consider, for example, a survey about whether the term “memory stick” is perceived to be a trade-mark or a generic term. A standard Teflon-test²⁰ survey approach *via* telephone may be employed. The name of the survey format originates with its first use²¹ in connection with testing the possible genericness of the mark “Teflon.” (With the support of survey evidence, its status as a trade-mark survived.) The Teflon-test format begins by defining “brand name” and “descriptive name” to survey participants, obtaining assurance that participants understand the distinction, and then asking participants to classify various terms in the product category — including, in this instance, the term “memory stick.” Such a survey produces three percentages, one for people who say “it’s a brand” (trade-mark), one for people who say, “it’s descriptive” (generic), and one for people who say “I can’t say” (no opinion). “Can’t say” could also mean, “I don’t know,” or “I don’t want to answer, because I’m afraid I’ll look stupid,” or “I can’t answer the question in the way you have asked it” or “I haven’t thought about it, and I can’t be bothered to think about it now.” The Teflon-test has received broad approval as a method for testing distinctiveness or genericism; but, like all structured tests, it depends on a particular conceptual model for tapping into consumer perceptions.

Another method for testing market reactions to the term “memory stick” is a mystery-shop of technology stores. Although such evidence has routinely been collected by private investigators at individual locations, a mystery shopping study can collect such evidence on a broader scale, from a statistically representative sample of retail locations. The use of rigorous statistical sampling, combined with quality controls on consistency and objectivity, allow mystery shopping data to meet the

¹⁹ *I.e.*, that non-response is also random.

²⁰ The Teflon test is the name of an established template for measuring the level of perceived brand significance in a given name. It is based on the evidentiary test in the early case preserving TEFLON’s trademark registration as a distinctive mark, in *DuPont de Nemours & Co. v. Yoshida Int’l*, 393 F. Supp. 502 at 526 (E.D.N.Y. 1975).

²¹ *DuPont*, *ibid.* at 525–527.

standards of social science expert evidence.²² In the case at hand, interviewers posing as shoppers may ask pertinent questions of clerks, or listen in on customer conversations in the store; information yielded by such investigations may reveal whether, in the ordinary course of trade, “memory stick” is used in a trade-mark sense or in a generic sense.

Social media and blog content may yield yet other information, not necessarily more truthful of what people think, but informative of what consumers en masse are exposed to on the Internet. One amateur photography blog site, for example, has a writer asking, “How do I quickly upload 500 photos from a memory stick?”²³ Such instances of use in social media can be systematically and objectively classified, and then counted. Advertising, news articles and technology-advice books may show a pattern of historical public use of the term “memory stick” that contributes to one hypothesis more than another. All of these sources provide different, but relevant, clues as to whether “memory stick” is likely to be perceived by consumers as a brand name or as a generic term. Each may be subjected to valid quantitative analysis. None of them is complete evidence on its own. But if they all point to the same conclusion, each reinforces a tentative conclusion that any one of them may suggest.

3. CONVERGENT VALIDITY EXPANDS OPPORTUNITIES TO PERSUADE

When one research result confirms another, any one of them is said to have convergent validity. Convergent validity refers to a situation where more than one independent source of evidence leads to the same conclusion, and thereby reinforces the truth of the conclusion.²⁴ The principle is best described by the well-known English saying: “If it walks like a duck, quacks like a duck and has feathers like a duck, it must be a duck.”²⁵ Convergent validity is a compelling concept because it matches, intuitively, how people develop confidence in a conclusion in their everyday decision-making. We frequently reserve judgment until a certain threshold of confidence is reached, by collecting information from trusted friends, from store clerks, from product review services, or from websites. Judges are sub-

²² R.M. Corbin & S. Carnegie, “Mystery Shopping Raised to Scientific Evidence,” *Vue*, the official magazine of the Marketing Research and Intelligence Association (September 2009), at 26–29.

²³ <<http://www.google.com/support/forum/p/blogger/thread?tid=76b2a424ceb30292&hl=en>>, accessed February 6, 2012.

²⁴ Sometimes referred to as “external validity,” it is found routinely in texts discussing social science research methods, for example in I.B. Weiner *et al.*, *Handbook of Psychology: Research Methods in Psychology* (New York: John Wiley & Sons, 2003), at 433–434.

²⁵ Giving several historical examples, Wikipedia suggests this principle has risen to the status of a “test” used by people to come to a conclusion about what should be obvious. See <http://en.wikipedia.org/wiki/Duck_test>, accessed October 3, 2012. The Duck Test appears in other cognitive science applications, including D.C. Funder, *Personality Judgment, A realistic approach to Person Perception* (San Diego: Academic Press, 1999) at 88.

ject to the same kind of intuitive confidence-building strategies as are other consumers. One piece of evidence may be tentative, two more persuasive; as independent pieces of evidence pile up, all of them converging on the same result, they give a decision-maker more and more certainty in what he/she believes to be the right choice.

The pursuit of multiple social science measurements with convergent validity has been successful in two recent appellate court cases in Canada. In the dispute between Jaguar Cars Ltd. and Remo Imports Ltd. (the latter producing JAGUAR brand suitcases and backpacks),²⁶ Jaguar Cars Ltd. submitted a multi-source program of social science expert evidence, including survey evidence, brand-awareness trend data, an expert marketing opinion, and historical media publicity records. All of the sources converged on the same conclusion: that JAGUAR was a famous brand of car threatened by confusion with Remo's JAGUAR brand luggage. The judge rejected all of the data collection methods but one, leaving Jaguar Cars Ltd. with a successful outcome: Remo's trade-mark for JAGUAR luggage was ordered expunged. A similar situation occurred in the Supreme Court's *Masterpiece* decision.²⁷ In that case, Alavida Lifestyles Inc.'s registration of "MASTERPIECE LIVING" in association with retirement home services was challenged by Masterpiece Inc., who had been using the phrase "Masterpiece the Art of Living" prior to Alavida's first use. Masterpiece Inc. submitted survey evidence, an analysis of the distinctiveness of "Masterpiece the Art of Living" based on expert literature, and an assessment of the marketing opinion of Alavida's expert witness. All three sources converged on the same conclusion: that Masterpiece Inc.'s use of its established slogans was threatened by confusion with Alavida's services. The court rejected the survey portion of Masterpiece's expert evidence, but accepted the evidence of distinctiveness, and the challenges to the opinion of Alavida's marketing expert. Masterpiece Inc. was successful in its appeal to have Alavida's trade-mark expunged.

In these and other cases, it took a different form of social science evidence to persuade the court, although the survey results yielded what was eventually accepted as the "truth" about the marketplace. It was propitious for counsel to have had evidentiary alternatives in hand. They turned out to be an insurance policy against the risk of survey rejection.

A court's rejection of a survey, while reaching the same conclusion as the survey gave rise to, poses a logical anomaly. Whatever evidence the court relies upon provides *de facto* convergent validity to survey results which come to the same conclusion. However, it is entirely the court's discretion to mean: "We have not relied on the survey; had the survey been the only evidence, we would not have found as we did; had the survey come to an opposite conclusion, we would have ignored it." Given that the court's analysis is not entirely predictable in advance, even for well-conducted surveys, a party submitting a survey expands its chances for succeeding by buttressing a survey with complementary social science evidence. This occurred to the benefit of the successful party in both *Jaguar* and *Masterpiece*.

²⁶ *Remo Imports Ltd. v. Jaguar Cars Ltd.* (2007), 2007 FCA 258.

²⁷ *Masterpiece Inc. v. Alavida Lifestyles Inc.*, 2011 SCC 27, [2011] 2 S.C.R. 387.

4. EVEN WHEN NOT STATISTICALLY RELIABLE, SURVEYS CAN BE VALID, OR VALID WITHIN LIMITS

Validity and reliability are the two main standards guiding social science, and have been adopted by the courts for evaluating survey evidence. Definitions of these terms, as used in social science and articulated in *Mattel*,²⁸ were given earlier. They represent distinctly different qualities of a measurement instrument, although courts have not always been as precise about the distinction as have social scientists.²⁹ Relevance is another criterion for assessing the usefulness of survey evidence, but is usually left to a court to determine.³⁰

In observing the limits on perfect statistical reliability, it is of some comfort and importance to remember that surveys which are valid but not demonstrably reliable are not without merit. Any valid survey provides truth about *some people*, even when those people are not representative of the relevant population.³¹ The extent to which the results may be generalized to *other people* in the relevant population is then a matter of judgment. For example, if survey respondents are demographically proportional to the pertinent population (by, say, age, gender, and geographical location), and have no vested interest in the content of the survey, then one may find it “reasonable” to generalize the results more broadly. Or at least, a decision-maker has the basis to ask, “Can I think of a plausible reason why the opinions expressed in this survey would be unrepresentative? Does the group who has been surveyed at least represent a materially important segment of the pertinent population? Do the results provide, at least, an illustration of what the submitter claims to be common opinion, to help challenge my own views of what I think is common sense?”

A survey sponsor can assist the cause of representativeness and reliability with best practice quality controls. These may include, but are not limited to:

- For mall surveys:

²⁸ *Mattel U.S.A. Inc. v. 3894207 Canada Inc.*, [2006] 1 S.C.R. 772, 2006 SCC 22.

²⁹ In *Masterpiece*, *supra* for example, the court criticized the survey design on grounds of validity, and subsequently summarized it, without support, as lacking reliability at para. 96:

For a survey to be valid, it seems elementary that there must be some consumers who could have an imperfect recollection of the first mark. Simulating an “imperfect recollection” through a series of lead-up questions to consumers will rarely be seen as reliable and valid.

The advisory note of the Marketing Research and Intelligence Society, *supra* note 8, reinforces the importance of the distinction.

³⁰ In Canada, the Supreme Court decision in *Mattel U.S.A. Inc. v. 3894207 Canada Inc.*, [2006] 1 S.C.R. 772, 2006 SCC 22 was the first court to clearly articulate the three standards of reliability, validity and relevance. Since relevance refers to the pertinence of the evidence to the matter to be decided, it is a matter for the court’s prerogative.

³¹ Expressed even with seemingly-maximum pessimism, one psychology textbook explains: “A survey using a biased, nonrepresentative sample is not necessarily worthless or uninteresting, but its results may not hold true for other groups.” In C. Wade *et al.*, *Psychology: Canadian Edition* (Toronto: Prentice-Hall, 2004), at 43.

— Setting demographic quotas for intercepting shoppers, according to the demographics of the local region.³²

— Qualifying people to participate according to characteristics discovered by a separate trustworthy source (*e.g.*, industry records, or a national telephone omnibus from a reputable firm.)

- For telephone surveys:
 - Applying a birthday method³³ for selecting a respondent within a particular household.
 - Including a portion of cellphone numbers, given the increase in cellphone-only households — conservatively estimated at 13 per cent of all Canadian households, and 50 per cent of households headed by younger adults in the 18–34 year old age bracket.³⁴
- For Internet surveys:
 - Requiring telephone validation of respondent identity and personal involvement in the interview; the number of interviews validated should match the validation standard of telephone interviews.
 - Using unique URL links for each respondent to prevent duplicate participation.
 - Including questions which test correspondence to demographic data given at the time the participant first signed up with the Internet panel.

The fact that valid surveys can deliver benefit without being, to a purist's standards, statistically reliable, helps to bolster the attractiveness of a convergent validity approach. A survey becomes one piece of a puzzle, a potentially valid piece, reporting on the views of a well-defined subset of the overall population. The other pieces of the marketplace picture remain to be assembled. Other pieces (if consistent in their outcomes) will broaden the representativeness of the result, until reliability can be inferred as a matter of reasonable judgment.

5. OTHER EVIDENTIARY OPTIONS

Convergent validity entails employment of multiple forms of social science evidence. These may include, but are not limited to, any of the following:

- In-person, telephone and Internet surveys (snail-mail surveys are no longer plausible)

³² The effect of such quota setting is to enable a demographically representative sample of the region of interest, thereby compensating for the unrepresentativeness which may characterize the shopping mall population.

³³ Asking to speak to the person whose birthday comes next in the household is an established method of choosing a random participant from any given household, avoiding possible bias of including only those who answer the phone most frequently.

³⁴ As reported by Statistics Canada in December 2010, accessible online at <<http://www.statcan.gc.ca/daily-quotidien/110405/dq110405a-eng.htm>>, accessed October 3, 2012.

- Mystery-shopping to obtain opinions and point-of-sale comments of store employees
- Social media monitoring, analyzed with statistical rigour
- Split-sample reliability tests of large scale data bases, analyzing whether results are replicated over different conditions
- Investigative Internet searches of blogs, articles, consumer advice sites
- Expert analysis of promotional strategies (including Internet promotions) to assess overlapping channels of trade³⁵
- Media archive content analysis, including foreign media advertising spilling over the relevant country's borders
- Market trend analysis obtainable from third-party information suppliers (such as Print Measurement Bureau, Bureau of Broadcasting Measurement, or industry association research services)
- Expert marketing literature review on topics of consumer attitudes, belief formation, perceptions, decision-making, or buying intentions

Each of the above approaches to evidence collection can be governed by the “scientific method”: null hypothesis, hypothesis testing, quantitative analysis, and statistical inference. Not all forms of social science research will be appropriate for each case. Fact situation, necessity and cost will determine the appropriate methodology. However, using at least two research methodologies, administered under strict quality controls, offers up the prospects of convergent validity. Each method has the chance of shoring up the other.

6. THE SPECIAL CASE OF SOCIAL MEDIA

“Without a doubt,” writes an analyst in the Law Gazette of Singapore, “one of the most significant implications of social media is how it has the potential to shape litigation strategy by introducing new sources of information.”³⁶ Social media and the Internet are a growing source of information for consumers³⁷ and have also been touted as “[having] given consumers around the world the most powerful voice they’ve ever had.”³⁸ Consumers taking the opportunity to make their voice heard through social media provide a data source for expert evidence. It is a matter

³⁵ See, e.g., *Quia Corporation v. Mattel, Inc. and Fisher-Price Inc.*, [2011] U.S.D.C. for the Northern District of California San Jose Division, Case Number C 10-1902 JF (HRL) at III A.: “The Court agrees that evidence of the parties’ respective internet marketing strategies may be relevant to the question of whether the two products are sold through convergent marketing channels that increase the likelihood of confusion.” Additional case citations are given therein.

³⁶ S. Low, “Lawyers and Social Media: A New Evidentiary Landscape,” in *Law Gazette*, an official publication of the Law Society of Singapore, 2011, published electronically at <www.lawgazette.com.sg/2011-06/132.htm>, accessed October 3, 2012.

³⁷ J.S. McKeown, *Brand Management in Canadian Law* (Toronto: Carswell, 2010) at 61.

³⁸ D. Kerpen, *Likeable Social Media: How to Delight Your Customers, Create an Irresistible Brand, and Be Generally Amazing on Facebook (and other social networks)* (New York: McGraw-Hill, 2011) at 4.

of context and judgment as to whether statements in social media are valid indicators of what the originators believe. On one hand, “this information may reflect candid and unbiased feedback which can be used like any other market research.”³⁹ On the other, individuals may exaggerate, advocate or invent for effect. But one can say for certain that social media statements provide *de facto* evidence of what millions of Internet users are exposed to.

Some encouragement for social media evidence exists in case law already. In *Chipotle Mexican Grill, Inc. v. Chipotles Grill of Jonesboro, Inc.*,⁴⁰ the United States District Court granted a preliminary injunction in part because of the plaintiff’s evidence from social media. The marks at issue being almost identical was “not enough to establish intent to confuse,”⁴¹ the court said, but the plaintiff’s evidence of actual confusion was “substantial.”⁴² That evidence consisted of e-mails sent to the plaintiff’s customer service website and evidence from two restaurant review websites. There was no survey involved.

[The restaurant forum webpages] reflect confusion as to the ownership of defendants’ restaurants. One website, www.urbanspoon.com, even lists plaintiff’s website as the website of defendants’ restaurants. A restaurant review on www.associatedcontent.com also seems to link defendants to plaintiff. There is actual confusion . . . Even if potential customers exercise reasonable care, it would be difficult to eliminate the confusion. The type of products marketed by plaintiff and defendants are similar; they are both Mexican grills. Plaintiff maintains that the price of a meal at its restaurants and defendants’ restaurants is comparable. Further, the e-mails and websites attached to plaintiff’s motion demonstrate that even after inquiry to outside sources, potential customers are still confused.⁴³

The *Chipotle* case and others⁴⁴ demonstrate reasonable prospects for social media evidence to be accepted into the fold of convergent validity. Finding compelling instances of actual confusion on the Internet may not require an expert, but analyzing the total content of one or more social media for statistical evidence of a conclusion is ripe for social scientific treatment.⁴⁵ Indeed, a social scientific analysis in the hands of a defendant may be equally valuable, should it suspect the plaintiff of cherry-picking a few perverse examples of confusion from the massive number of social media entries on the Internet.

Quality controls are under continuous development to address data quality is-

³⁹ *Supra* note 26 at 309.

⁴⁰ 2011 U.S. Dist. LEXIS 61678 (E.D. Ark., Western Division 2011).

⁴¹ *Ibid.* at section 3(c), para. 8.

⁴² *Ibid.* at section 3(c), para. 9.

⁴³ *Ibid.* at section 3(c), paras. 9 and 10.

⁴⁴ *E.g.*, in Canada, *Candrug Health Solutions Inc. v. Thorkelson*, 2007 FC 411.

⁴⁵ A search on the Canadian Legal Information Institute <www.canlii.org> on October 3, 2012 found no references to social media used as evidence in Canadian intellectual property cases.

sues in software specifically designed for social media monitoring.⁴⁶ Standards are particularly important in distinguishing between casual observation of what some people type, and statistically defensible expert evidence. The level of social scientific integrity will be determined by the ability to transform apparently qualitative information into quantitative data, through objective content analysis, made amenable to statistical analysis. Available standards include many of those reminiscent of survey data: preserving the dates of data collection, maintaining records through screen-capture of relevant web-pages, sampling across different search engines, across different types of websites (blog sites, video sites, news sites), weighting, using scientific content analysis to interpret entries or choosing objective search terms to “question” search engines.

Social media have already been widely touted as sources of litigation evidence. The opportunity for defensible expert evidence is to reach beyond selective examples from a stunningly large number of possibilities, to a statistically analyzable and valid snapshot of the entire landscape of electronic messages — much as parades of *viva voce* witnesses came to be efficiently replaceable by well-conducted surveys.

7. CONGRUENCE WITH THE FOUNDATIONS OF BUSINESS, SCIENCE AND TRADE-MARK LAW

A convergent validity approach to expert evidence is consistent with parallel trends in 21st century knowledge development. “Integrative thinking” is promoted by many sources as a path to greater business success. In his book, *The Opposable Mind*,⁴⁷ author Roger Martin explores the superior performance of integrative thinkers, those who see relevance in diverse sources of information. Although individuals are rarely conscious of the cognitive steps taken when making a decision, integrative thinkers are distinguished by how they take those steps. They take a broader view than others of what could be salient to their decision.

They welcome complexity because they know the best answers arise from complexity. And they feel confident that they will not get lost along the way but emerge on the other side of the problem with a clear resolution ... Integrative thinking produces possibilities, solutions, and new ideas.⁴⁸

The effect of good advocacy is to put a decision-maker in the frame of mind where he/she views a collection of evidence as relevant, logically assembled, and supportive of one outcome in preference to another. Social science evidence with convergent validity provides the objective underpinnings to facilitate such a frame of mind on the part of a decision-maker. It expands the landscape of relevant, trustworthy, information sources, integrates them into a logical perspective of the mar-

⁴⁶ The authors find data quality issues to be a frequent topic in the business media, as research companies describe their continuous strategies of improvement. See, e.g., A. Pettit, “Like a Survey,” *Vue* (October 2011) at 13.

⁴⁷ R.L. Martin, *The Opposable Mind: How Successful Leaders Win Through Integrative Thinking* (Boston: Harvard Business School Press, 2007) 25-26.

⁴⁸ *Ibid.* pp. 41, 48.

ketplace, and demonstrates where all roads seem to lead.

As a scientific concept, convergent validity can also rely on the technical underpinnings of Bayesian statistics. The model of gathering multiple sources of independent evidence, combining it with prior knowledge, to reach a sound decision, is the basis of the Bayesian approach.⁴⁹ Bayes' Theorem provides a logical, quantitative, science-compatible framework for information integration. This statistical grounding reinforces its scientific legitimacy.

Finally, a convergent validity approach to knowledge advancement also has the professional support of the worldwide marketing research profession. Many marketing research firms now promote the benefits of multiple sources of research to gain a higher level of confidence in the results and create "a more complex and fulsome picture."⁵⁰ When asked why E-Rewards had brought a social media research firm into the fold of its traditional survey research business, its CEO replied in a publicized interview:

I wouldn't say it's a departure, it's a complementary thing. Stepping back and looking at what the E-Rewards group does, we are in the business of providing data to drive insights. As the research industry evolves, it's starting to think about the different sources of data it can tap into. There's life beyond survey research, we know that.⁵¹

Canada's national Marketing Research and Intelligence Association, once dominated by survey research professionals, now offers an annual "Best Integration Award" for research programs that incorporate multiple methodologies.

In summary, forums for the advancement of knowledge in business and science are more assertively advancing a multi-input approach. This fact argues favourably for adoption of such an approach to social science expert evidence, since congruence with methods of business, professional practice and science has traditionally been the courts' standard of acceptability.

Finally, the prospect of multiple converging sources is congruent with the statutory multi-factor test of trade-mark confusion.⁵² Case law has also contributed factors not explicitly laid out in statute. Similarity of the parties' channels of distribution, the strength of distinctiveness of the marks, the sophistication of prospective purchasers, the similarity of marketing strategies, and the extent of overlap in the parties' respective markets may all be factors relevant to the court's analysis. Expert evidence, which builds on more than one of these to yield a more holistic picture of the marketplace, delivers more potential value to a trier of fact. A survey may or may not be an essential component in that evidentiary program.⁵³

⁴⁹ Bayesian Statistics is a long-established subject in mathematics, now the subject of a dedicated organization described at <<http://bayesian.org>>, accessed October 3, 2012.

⁵⁰ Nik Nanos quoted in "In Conversation with Nik Nanos," *Vue* (December 2011) at 13.

⁵¹ Reported in an industry e-newsletter, at <<http://www.research-live.com/life-beyond-surveys/4005176.article>>, accessed October 3, 2012.

⁵² As laid out in s. 6 of the *Trade-marks Act*.

⁵³ D. Sarel & H. Marmorstein, "The Effect of Consumer Surveys and Actual Confusion Evidence in Trademark Litigation: an Empirical Assessment," in *Trademark Reporter* (2009), 99, at 1416-1422.

8. RECOMMENDATIONS ARISING

Decision-makers in business, governments and courts need not wring their hands about the declining representativeness of surveys. Surveys can be reliable within defined populations, as long as one can define the population from which they have been sampled. For example, a margin of error can be calculated based on the population of visitors to a tradeshow, or on the population of customers on a given list. Reasonable inference of reliability may continue to be made about mall intercepts and telephone surveys in the presence of superior representative sampling techniques. Response rate may interfere with reliability but, historically, that has always been the case.

Alternatively, surveys may have no quantitatively estimated reliability but still be valid within their sampling parameters. A valid survey gives a true picture of the views, perceptions or intentions of at least that group which has been surveyed. Then other factors may intervene to determine reliability. For example, psychology experiments on university students have been the basis of a good deal of our confirmed knowledge about how the brain is wired to make judgments.

Whatever reasonable judgment is employed to draw inferences beyond a survey sample, surveys do not necessarily tell the whole story, and may benefit from confirmation or elaboration through other valid forms of social science. Courts should look not to "survey evidence," but to social science evidence, part of which may be delivered through a survey. Experts should be clear and realistic about what a single survey proves. They have the opportunity to reinforce or clarify survey data through demonstration of convergent validity with other methods. There is a broad array of social scientific methods having the requisite characteristics of reliability and validity whose potential has not been fully exploited. In the absence of opposing evidence to the contrary, a collection of converging results is arguably compelling. Opponents who criticize a survey for the sake of criticizing (and there is always fodder for criticism) would be harder put to explain how a survey constructed differently would change the collective conclusion. If it walks like a duck, quacks like a duck and has feathers like a duck, there is little merit in challenging whether it can swim.

How does convergent validity address the challenges of a post-*Masterpiece* environment? The Supreme Court wrote that survey evidence would be unnecessary if it fell within the common sense or everyday experience of the trier of fact. Surveys are therefore more likely to be approved at the case management stage if a litigant can argue successfully that its results would hold some surprise for a court beyond what common sense or everyday experience would predict. But if a survey is allowed for the very reason that it would challenge a court's preconceptions, then it is destined to be part of an uphill battle of persuasion, with a heavy anticipated burden of proof. It may be a weighty burden for a single survey to carry alone.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.