



# CorbinPartners' Handbook on Preference Claim Substantiation

---

Advice on preparing comparative advertising claims using survey evidence  
First Release, October 2015

## Table of Contents

Similar Sounding Claims Can Have Important Differences .....	2
Reliability: Achieving a Defensible Level .....	3
Validity: Key Issues .....	5
Relevance: How to Support Specific Claims.....	6
Guidelines for Claim Communication .....	8

*Courtesy of*

**CorbinPartners Inc.**  
**Phone:** 1-416-413-7600  
**Web:** [www.corbinpartners.com](http://www.corbinpartners.com)  
**Email:** [info@corbinpartners.com](mailto:info@corbinpartners.com)

---

**NOTE:** This reference document was prepared in October 2015, and is subject to modification as new information or standards arise. For more information, please see the guidelines for comparative advertising from Advertising Standards Canada, or contact CorbinPartners.

---

## Similar Sounding Claims Can Have Important Differences

Preference claims require adequate and proper testing, and non-misleading communication. This brief guide provides summary advice on using survey evidence for claim substantiation and communication.

A company's great ginger ale product might be advertised as follows:

***“More Canadian ginger ale drinkers prefer the taste of Gin Fizzy to that of the leading national brand.”***

Or it might be advertised this way:

***“Gin Fizzy is Canada's most preferred ginger ale.”***

Did you spot the important differences? Surveys to substantiate those two claims are likely to have different test populations, different comparator products, and different survey questions.

Whatever survey design is chosen must be conducted according to scientific standards, with a statistically significant result. And there are many more specific test details to think about, in order to prepare for a possible face-off with an indignant competitor, who complains through a court or regulatory tribunal. Of course, no amount of thoroughness offers an iron-clad guarantee that evidence will be accepted by a court or regulator, but the guidelines below will at least ensure that your evidence will be scientifically defensible and supported by precedent.

It all boils down to three criteria of good science: “reliability” (reproducibility of the results), “validity” (the true capture of the objects of measurement) and “relevance” (results pertinent to the precise claim being made). Besides being consistent with good science, these principles incorporate guidance from the Supreme Court of

Canada for the evaluation of survey evidence.<sup>1</sup>

While we confidently advocate scientific standards, the results must be straightforward to describe in layperson terms — they have to make sense. Particularly when there is conflicting evidence between two parties, and when the opposing arguments get technical and obscure, judges are inclined to fall back on their common sense to choose between the two.



<sup>1</sup> *Mattel, Inc. v. 3894207 Canada Inc.* 2006 SCC 22

## Reliability: Achieving a Defensible Level

---

Reliability means reproducibility of results. Three key components of reliability are (a) a random, representative sample, (b) a sufficiently large sample size, (c) quality controls for consistency — to minimize non-sampling error.

### **(a) Obtaining a random, representative sample**

In today's consumer environment, it is near-impossible to generate a provably representative sample. This section offers guidelines for obtaining what can be an arguably or plausibly representative sample.

**Telephone surveys** can readily generate a random sample of households. But their response rates are quite low, usually less than 10%, and frequently lower for cell-phone-only households. Representativeness can be increased by diligent call-back strategies for busy, not-at-home and not-available households, by randomizing the person in the household who is chosen to respond, and by weighting the sample for age and gender within each region to be consistent with Census demographics. These strategies are all consistent with the best practices of the Marketing Research and Intelligence Association – MRIA ([www.mria-arim.ca](http://www.mria-arim.ca)).

**Mall surveys** can readily generate a random sample of visitors to any given mall, but the sample may not be representative of all shoppers in the trade area or in other trade areas. Representativeness can be increased by carrying out interviews in more than one mall and more than one region. Courts and regulators have been sympathetic to the significant costs of in-person interviews in geographically disparate locations, and have accepted evidence combined from a modest number of locations. The Appendix shows the number of malls or other interviewing locations that have proven acceptable to certain regulatory bodies and industry associations. The representativeness of mall surveys can also be enhanced by screening for Census-based quotas on age and gender for that geographic region, and for any other demographic variables known to be relevant to the product test. For variables where actual proportions are not known or knowable, it is reasonable to assume (with good randomized selection techniques) that the right proportions will fall out in the sample.

**Internet surveys** depend on volunteers who, by definition, are not representative of the broad consumer population. However, within the very large respondent base of Internet panels, a representative sample can be chosen and demographically balanced by age, gender and region to Census proportions. Some specialized audiences, like cigar-smokers, are most efficiently accessed through Internet contact (and survey costs may be prohibitive without access to an Internet population). The major quality control issue for Internet surveys is assuring the *bona fides* of those presenting themselves as respondents.

A final word on representativeness: with a scientifically-designed control group as part of the survey process, a perfectly representative sample is less important than would otherwise be the case. Where a control group is present, the interpretation of the survey results is usually based on a comparison between the test group and control group; if survey results are systematically replicated for relevant sub-groups within the sample, it can be inferred that any weighting of the sub-groups to the population at large will generate similar results. That is, the sample may “behave” like a representative sample, with respect to the main measurement of interest.

### **(b) Choosing a sample size—how many is enough?**

For a broad-based survey, Canadian precedents suggest a minimum sample size of 300-400 respondents.<sup>2</sup> That sample size range is exclusive of a standalone sample, if required, for a scientific control condition to measure “cause and effect”. Control condition results typically have low variation in the range of responses, and for technical reasons that means a smaller sample will suffice. A national sample of 100-120 can usually be defended for a control condition.

One more caveat: If geographic differences, age differences, or other segmented differences are important to establish, a larger overall sample may be required. Any single subset of data that is important for comparison purposes should have 30 to 50 individuals in it as a minimum.<sup>3</sup>

**✓ Consider a minimum sample size of 300 - 400, plus a standalone sample of at least 100 - 120 wherever a separate control condition is required.**

The Appendix contains guidelines from different authorities related to sample size and geographic distribution.

### **c) Reliability means tight control over consistency**

Researchers use the term non-sampling error to refer to the deviation from a “true” result that arises from factors other than imperfect sampling. Non-sampling error potentially includes such elements as variances in interviewer rapport with respondents, variations in how products are prepared or presented, lapses in respondent attention, and other factors that undermine perfect consistency of administration.

---

<sup>2</sup> See Advertising Standards Canada – Guidelines for the Use of Research and Survey Data to Support Comparative Advertising Claims, Section 4.6 (c), pg. 5. For case law see, e.g. *Kraft Suchard Jacobs (Schweiz) AG v. Hagemeyer Canada Inc.*, (1998) 78 C.P.R. (3d), Ontario Court (General Division), Court File no. 97-CV-130843.

<sup>3</sup> That is because the theory for calculating statistical reliability is based on minimum sample sizes of 30, selected randomly.

Reasonable control of non-sampling error is achieved in part, by following published standards of the market research industry.<sup>4</sup> The MRIA has assembled “Gold Seal” standards representing best practice, and regularly audits its Gold Seal members’ compliance with them. The following test components beyond sampling (where applicable) should be subjected to quality control and operational consistency.

- ▶ Staff Training
- ▶ Product Procurement
- ▶ Product Preparation
- ▶ Product Serving/Display
- ▶ Respondent Screening
- ▶ Interview Execution
- ▶ Participant Honoraria
- ▶ Interview Validation
- ▶ Data Preparation
- ▶ Data Verification
- ▶ Data Analysis

**✓ Non-sampling error must be controlled through rigorous consistency of administration.**

## Validity — Key Issues

---

Validity refers to the extent to which a survey captures people’s true views and measures them authentically.

One of the key issues in achieving validity is respecting the fact that some people truly have no preference. They should not be forced to choose between two options where they have no basis for preference. They should have access to an option to say “I have no preference” or “I don’t know” or “I have no opinion”.<sup>5</sup>

What should you do with the “no preference” votes in assessing strength of preference for your product? Proceed with caution! After all, results that show a split of one-third / one-third / one-third among “Prefer A / Prefer B / No preference” are indistinguishable from random responding. Further, if the percentage of “no preference” answers forces the preference for your product to be less than 50%, then a claim of “majority preference” is no longer available.

Other key issues in the achievement of validity include the design of questions which are neither leading nor loaded, and the design of a control condition (where needed) that is neutral with respect to its possible influence on results.

<sup>4</sup> MRIA 2015 Standards - <http://mria-arim.ca/about-mria/standards/code-of-conduct-for-members>

<sup>5</sup> Advertising Standards Canada – Guidelines for the Use of Research and Survey Data to Support Comparative Advertising Claims, Section 4.6 (c), pg. 7

## Relevance: How to Support Specific Claims

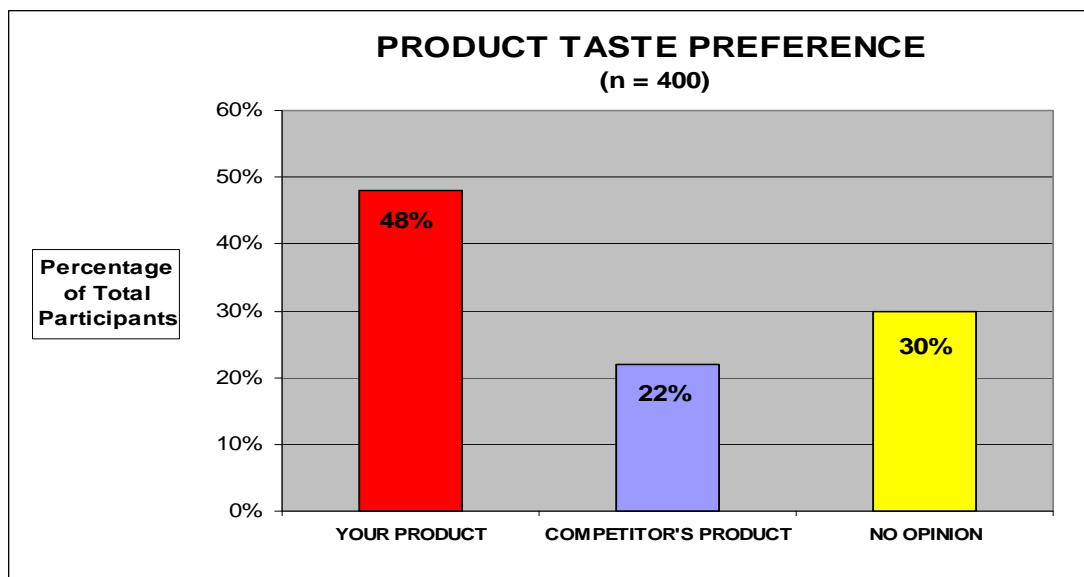
Relevance refers to the pertinence of the survey results to the perceived meaning of the claim. The guidelines on relevance, below, assume a national 2-product claim test conducted with 400 participants across 4 markets, supported by a test that has adhered to scientific standards for reliability and validity.

### A “majority prefers” claim

When testing your product against a single competitor product, it is tempting to treat any level of preference above 50% as sufficient to make a claim that “the majority prefer...”. However, don’t forget about the sampling and non-sampling error. The margin of error for a random national sample of 400 is just under 5%. Conservatively assuming non-sampling error to be approximately equal to the sampling error, a “majority” claim should only comfortably be made when measured preference of your product is over 60% (that is, 50% + 5% sampling error + 5% non-sampling error).

### “Our product is preferred over Competitor X” claim

Even without a clear majority preference, a claim of preference can be made whenever there is a statistically significant difference between preference for your product and preference for the competitor’s product. This could occur, for example, with the following range of preference percentages:



In other words, among those with a preference, your product is the clear statistical winner, but the “No preference” answers remove the option to make a “majority prefers” claim. Then you could still make a claim such as:

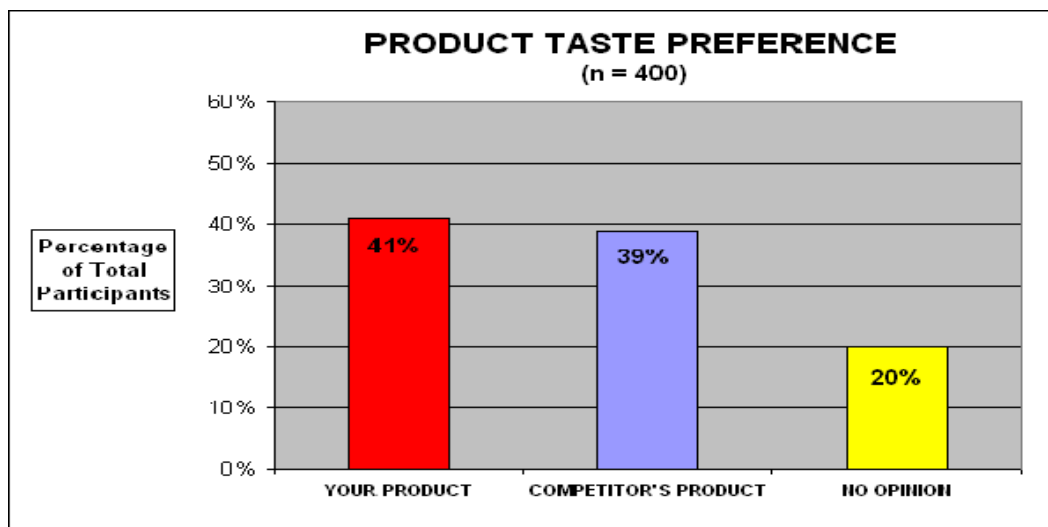
***“In a cross-country preference test between Gin Fizzy and the leading national brand, more ginger ale consumers preferred the taste of Gin Fizzy.”***

Or....

***“Of ginger ale consumers stating a preference in a cross-country taste test, more preferred Gin Fizzy to the leading national brand.”***

### **A “just as much” claim**

A parity situation applies when the percentage preference for your product is statistically the same (within the margin of error) as that of your competitor’s product. This could occur, for example, with the following range and percentage of preferences:



In this case, the claim can report on parity:

***“Canadian ginger ale consumers say they like the taste of Gin Fizzy just as often as the leading national brand.”***

Or...

***“In a cross-country taste test, ginger ale consumers preferred Gin Fizzy as often as the leading national brand.”***

A final word: the claim must always refer to, or unambiguously imply, the pertinent population from whom support was gathered. In all of the cases above, the audience in question was ginger ale consumers and the claims referred only to them.

## Guidelines for Claim Communication

---

According to the *Canadian Code of Advertising Standards*, all advertising should be conducted based on standards of honesty, truth, accuracy, fairness and propriety.<sup>6</sup> Here are some principles for interpreting that directive:

**Lifetime of research data:** Data must be reliable and valid at all times during which research-based comparative advertising claims are made. This requires that the relevant market conditions should be reasonably comparable to those when the data were collected.<sup>7</sup> A 12-month timeframe is a good rule of thumb for data currency, assuming no changes in the product or competitive environment that would materially impact the accuracy of the claim. A material change might include: a change in product formula, or a change in the competitive brand being compared, or a new discovery that changes the way consumers think about the product category.

The 12-month timeframe also assumes there are no seasonality factors that would affect product preference.

✓ ***Assuming no significant changes to the competitive landscape, consider refreshing research evidence used to support a marketing claim after no more than 12 months, and sooner if market conditions change in a material way.***

**Disclaimers:** The marketing claim should be capable of standing on its own, and not be misleading or incomplete when read without a disclaimer.

Disclaimers should not provide any contradictory information to the featured claim, and should be clearly visible or audible. The latter will pre-empt opportunity for opponents to claim that consumers were misled by virtue of not having processed the information in a disclaimer.



---

<sup>6</sup> Advertising Standards Canada – The Canadian Code of Advertising Standards - <http://www.adstandards.com/en/standards/canCodeOfAdStandards.aspx>

<sup>7</sup> Advertising Standards Canada – Guidelines for the Use of Research and Survey Data to Support Comparative Advertising Claims, Section 4.6 (c), pg. 9



It is advisable, though not always essential, that the bases for a claim be disclosed in a footnote by referring in summary form to the supporting test results.

✓ ***Disclaimers may be used to clarify the basis for a claim, but should not be used to complete the meaning of the claim itself.***

**Underperformance in specific segments:** A question arises as to the defensibility of a claim where a particular demographic group may show results different from the overall population. For example, there may be a preference for your product at the national level, but not so in Quebec or Manitoba. It may still be acceptable to make a statement about preference in the Canadian population, however, as long as there is low risk that residents of Quebec or Manitoba (in this example) would rely on the assumption that claim is true for their home province.

✓ ***If some demographic group fails to show a preference, assess the risk and consider appropriate claim wording to specify the group(s) or region(s) to which the claim applies.***

- APPENDIX -

**GUIDELINES FOR SURVEY SAMPLE SIZE AND DISTRIBUTION**

SOURCE	SAMPLE SIZE	CONFIDENCE LEVEL	REGIONS
<p align="center"><b>Report on Best Practices in Conducting Surveys In Trademark Matters - 2013</b></p> <p align="center"><b>(INTA – Courts and Tribunals Subcommittee)</b></p> <p><a href="http://www.inta.org/PDF%20Library/INTA%20Report%20on%20Best%20Practices%20in%20Conducting%20Surveys%20in%20Trademark%20Matters.pdf">http://www.inta.org/PDF%20Library/INTA%20Report%20on%20Best%20Practices%20in%20Conducting%20Surveys%20in%20Trademark%20Matters.pdf</a></p>	Minimum 300	No reference	Appropriate representation
<p align="center"><b>Advertising Substantiation – 2015</b></p> <p align="center"><b>(Marketing Research Association – MRA)</b></p> <p><a href="http://www.marketingresearch.org/issues-policies/best-practice/advertising-substantiation-and-standards-conducting-research">http://www.marketingresearch.org/issues-policies/best-practice/advertising-substantiation-and-standards-conducting-research</a></p>	Minimum 300 (min. 100 for subgroup / control group)	95%	No reference
<p align="center"><b>Reference Manual on Scientific Evidence-2011</b></p> <p align="center"><b>(Federal Judicial Center/National Research Counsel)</b></p> <p><a href="http://www.fjc.gov/public/pdf.nsf/lookup/SciMan3D01.pdf/\$file/SciMan3D01.pdf">http://www.fjc.gov/public/pdf.nsf/lookup/SciMan3D01.pdf/\$file/SciMan3D01.pdf</a> (Pg. 381)</p>	No reference	95%	No reference
<p align="center"><b>Advertising Guidelines – 2014 (NBC)</b></p> <p><a href="http://nbcuadstandards.com/files/NBC_Advertising_Guidelines.pdf">http://nbcuadstandards.com/files/NBC_Advertising_Guidelines.pdf</a> (Pg. 25-26)</p>	Minimum 300	95%	At least 4 regions; At least 2 markets per region
<p align="center"><b>Guidelines for the Use of Research and Survey Data to Support Comparative Advertising Claims – 2012</b></p> <p align="center"><b>(Advertising Standards Canada)</b></p> <p><a href="http://www.adstandards.com/en/ASCLibrary/guidelinesCompAdvertising-en.pdf">http://www.adstandards.com/en/ASCLibrary/guidelinesCompAdvertising-en.pdf</a> (Pgs. 5-6)</p>	Minimum 300	95%	At least 4 regions

SOURCE	SAMPLE SIZE	CONFIDENCE LEVEL	REGIONS
<p><b>Advertising Standards and Guidelines – 2014 (ABC)</b></p> <p><a href="http://abcallaccess.com/wp-content/uploads/2014/07/2014-Advertising-Guidelines-.pdf">http://abcallaccess.com/wp-content/uploads/2014/07/2014-Advertising-Guidelines-.pdf</a> (Pg. 22-25)</p>	Minimum 200 (for superiority claim)	95%	At least 4 regions
<p><b>What IP Attorneys Should Know About Expectations and Costs for Survey Research (The TASA Group – Technical Advisory Service for Attorneys)</b></p> <p><a href="http://www.tasanet.com/knowledgeCenterDetails.aspx?docTypeID=1&amp;docCatID=13&amp;docID=43">http://www.tasanet.com/knowledgeCenterDetails.aspx?docTypeID=1&amp;docCatID=13&amp;docID=43</a></p>	Minimum 200-300	95%	At least 4 regions
<p><b>Legal Review : 9 out of 10 regulators prefer substantiated preference claims – 2013 (Venable LLP)</b></p> <p><a href="http://www.lexology.com/library/detail.aspx?g=62243951-bc78-40ea-8940-b564b2b972b0">http://www.lexology.com/library/detail.aspx?g=62243951-bc78-40ea-8940-b564b2b972b0</a></p>	Minimum 250	No reference	At least 4 regions
<p><b>Conducting Community Surveys - A Practical Guide for Law Enforcement Agencies -1999 (US Department of Justice -Bureau of Justice Statistics)</b></p> <p><a href="http://www.bjs.gov/content/pub/pdf/ccspglea.pdf">http://www.bjs.gov/content/pub/pdf/ccspglea.pdf</a> (Pg. 14)</p>	Minimum 200-250	No reference	No reference
<p><b>Survey Methods and Practices – 2010 (Statistics Canada)</b></p> <p><a href="http://www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf">http://www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf</a> (Pg. 152-155)</p>	Minimum 400	95%	No reference

SOURCE	SAMPLE SIZE	CONFIDENCE LEVEL	REGIONS
<p><b>Case in ASC Reviews – 2014 decision by National Advertising Division (NAD), administered by Counsel of Better Business Bureaus</b></p> <p><a href="http://www.ascreviews.org/2014/11/nad-recommends-mom-brands-discontinue-taste-test-claims-challenged-by-post-advertiser-to-appeal/">http://www.ascreviews.org/2014/11/nad-recommends-mom-brands-discontinue-taste-test-claims-challenged-by-post-advertiser-to-appeal/</a></p>	No reference	No reference	Minimum of 2 markets in each of 4 regions
<p><b>Standard Guide for Sensory Claim Substantiation - 1998 (American Society for Testing and Materials)</b></p> <p><a href="#">%20ASTM%20Standards/ASTM%202004%20PART%20III/ASTM%202004%20Volume%2015.08%20Sensory%20Evaluation_%20Livestock,%20Meat,%20and%20Poultry%20Evaluation%20Systems_OTHERS/E%201958%20-%2098%20%20_RTE5NTG_.pdf</a></p>	Minimum 300	95%	Minimum of 2 markets in each of 4 regions

**NOTE:** The sample size / distribution guidelines presented above are based on large and diverse populations. For very small or specialized populations, smaller samples may suffice. If analysis of sub-samples is required (e.g., different age groups, particular geographic regions, control group), the overall sample may need to be larger.

For further information on sample size, see Corbin, R., “*Sample Size: how big is big enough?*” adapted from Chapter 9 of the Canadian Trademark Law Benchbook, Toronto: Carswell, 2014; copy available from CorbinPartners Inc.