Antitrust, Vol. 31, No. 3, Summer 2017.

Claims Substantiation for New Technologies

BY CHRISTOPHER COLE

IT IS A FUNDAMENTAL REQUIREMENT OF advertising law that factual claims for products and services must be substantiated. In its 1972 decision in *Pfizer, Inc.*,[1] the Federal Trade Commission established the principle that an advertiser must possess and rely on a "reasonable basis" to substantiate advertising claims. This was reiterated in the FTC's Advertising Substantiation Policy Statement of 1983, which stated: "Objective claims for products or services represent explicitly or by implication that the advertiser has a reasonable basis supporting these claims."[2]

In a series of decisions and guidance dating back several decades, the FTC has elaborated on the kinds of support required to constitute a reasonable basis for ad claims. In the case of health and technical or science-based claims, the FTC has required the advertiser to possess "competent and reliable scientific evidence."[3] For example, in *Mars Petcare US, Inc.*, the Commission's Order required the respondent to possess "competent and reliable scientific evidence" before communicating any claims regarding the effect of pet food on the longevity of pets.[4] The Commission defined "competent and reliable scientific evidence" as tests, analyses, research, or studies that have been conducted and evaluated in an objective manner by qualified persons and are generally accepted in the profession to yield accurate and reliable results." On occasion, the Commission has elaborated on the kinds of tests it would consider to be competent and reliable in order to provide greater specificity and eliminate enforcement ambiguity.[5]

For many products and product categories, standards for substantiation have been built over decades of experience. These standards are often marshaled in the trade association context, or by private standard-setting bodies, such as the American Society for Testing and Materials (ASTM) or International Electrotechnical Commission (IEC). Through a consensus process, and often tested through the litigation process, standards for product testing are refined over time until they reach widespread acceptance. Decision making bodies like courts and the National Advertising Division (NAD) of the Council of Better Business Bureaus tend to give great weight to consensus standards where they exist and are a good match to the claims being made.

Christopher Cole is a partner of Crowell & Moring. He represented parties in the Clairol, Comcast, Dyson, Kimberly-Clark, Procter & Gamble, Miller Brewing, and Zero Technologies cases discussed in the text.

But the situation is murkier where no such consensus or industry standards exist, as is usually the case for novel products and claims. There, the advertiser seeking to communicate regarding the newly developed benefit must prove the benefit through test methods that are typically not widely known or accepted. Such tests may borrow from existing standards, but frequently lack the recognition or consensus approval inherent in industry standard tests.

This article discusses the hallmarks of "good" testing for novel products and claims and sets forth general considerations for substantiation of claims regarding innovative products and novel features. As discussed below, FTC precedent and guidance suggests more restrictive policy for substantiating novel claims involving foods, drugs, and medical devices, which often implicate health and safety, than for other kinds of products. This certainly makes sense from a policy perspective.

Sources of Precedent for Testing Standards

The FTC's Influence. The FTC's *Pfizer* decision set the baseline considerations for evaluating whether an advertiser has "competent and reliable scientific evidence." What kinds of testing are required to support the claim will depend foremost on the type of claim being asserted. The FTC has elaborated on this requirement in various published orders, listing, for example: "tests, analyses, research studies, or other evidence based on the expertise of professionals in the relevant area, that have been conducted and evaluated in an objective manner by persons qualified to do so, using procedures generally accepted in the profession to yield accurate and reliable results."[6] In *Pfizer* itself, the FTC discussed five factors relevant to the quantum of evidence required to support a claim, as follows:

(1) the type and specificity of the claim made—e.g., safety, efficacy, dietary, health, medical;
(2) the type of product—e.g., food, drug, potentially hazardous consumer product, other consumer product;
(3) the possible consequences of a false claim—e.g., personal injury, property damage;
(4) the degree of reliance by consumers on the claims; and
(5) the type, and accessibility, of evidence adequate to form a reasonable basis for making the particular claims.[7]

Thus, the FTC has articulated clearly that claims significantly impacting human health will require the highest level

of support, but beyond this sliding scale, the decision does not provide much of a roadmap for specific testing requirements.

The FTC does repeatedly emphasize that the opinions of experts in the field matter greatly, and while unanimity of opinion in support of a claim is not necessarily mandated, some of its decisions (particularly those earlier this decade) tended to reject claims that were not supported by multiple studies and published research, including a majority of relevant experts in the field.[8]

Such standards can be unforgiving on new and novel products, as well as on emerging science. They reflect a policy judgment that disfavors the advertising of novel attributes, particularly relating to food and health, where Factors 2 and 3 of the *Pfizer* factors are implicated.[9] However, they are less restrictive for novel technologies not impacting health or safety.

NAD's Role. NAD has largely followed the FTC's lead in the evaluation of claim substantiation. While not precisely articulating its reliance on *Pfizer*, NAD typically applies close scrutiny to claims support, and in a series of published decisions going back several decades, has broadly adhered to principles requiring sound consumer research, good laboratory practice, and other recognized hallmarks of good scientific methods, such as blinding, randomization, proper sampling techniques, and use of controls.

In many respects, NAD decisions are more informative and comprehensive than those available from the FTC. NAD publishes with each decision an opinion setting forth the arguments and a summary of evidence, and it includes an explanation of its rationale. FTC standards are often embedded in the form of orders and complaints, with little explanation for why certain tests were found to be reliable or unreliable.

Industry Standards. Many industries have adopted standard testing methods that provide a recognized baseline on which companies can base claims. For example, in the water filtration area, NSF International has published a series of widely accepted water testing methods[10] for various contaminants that are recognized as the industry standard by which laboratories can test filtration performance and compare results across labs.[11] While NSF itself conducts such tests on a paid basis at the request of industry, its standards are available for reference by other labs, which can also conduct the tests.

ASTM has also developed numerous test methods, ranging from tests on biodegradation (ASTM D5511-11) to a Guide for Substantiation of Sensory Claims (ASTM E1958-16).[12] The ASTM test methods have been developed using a consensus process, with leading experts from industry, consulting, and academia collaborating in an iterative process to come to a consensus set of methods and recommendations. The process is generally transparent, with draft standards made available to interested parties for review and comment.

Other widely used testing organizations include the IEC, which is the European equivalent of ASTM, the American National Standards Institute (ANSI), International Standards Organization (ISO), National Fire Protection Association (NFPA), a variety of federal government entities and numerous trade associations.

In recent years, the number of private standard-setting and certification organizations has proliferated. The FTC has expressed concerns about this proliferation in the area of "green claims," noting in conjunction with revisions to the Green Guides that some organizations may lack the credibility, expertise, or independence to publish and certify compliance with standards.[13] Indeed, the FTC has brought enforcement against certifying organizations alleged to have "rubber stamped" environmental claims, and warned others regarding the perils of self-labeling with seal claims that are unsupported.[14]

The better consensus standards are not only developed with input from the leading experts in the field, they are often the subject of round-robin testing to ensure that results are repeatable across different test centers, as discussed below. In addition, many standards have been evaluated in terms of their correlation with "real world" measurements. Such tests attempt to determine whether the laboratory or idealized test conditions predict real-world results with reasonable precision. ASTM and other entities sometimes publish these correlation studies in appendices or other background to the main test standard.[15]

The Courts. Although court opinions can, in the right cases, set forth helpful commentary and precedent for testing standards, and in certain industries have laid down clear markers for test requirements, they are at best an incomplete reference source when it comes to advertising substantiation. As an initial matter, the courts are not experts in technical issues of advertising substantiation, but are required to rule on the facts before them. Judges' decisions sometimes turn on issues of credibility (e.g., was the chief scientist a believable witness) as much or more than science. They do not have the perspective that the FTC and NAD have, which is honed by reviewing thousands of ads and underlying tests over many years.

That said, the best-developed area of law in judicial opinions under the federal false advertising statute, known as the Lanham Act,[16] tends to be in the evaluation of consumer surveys and copy tests, where a long line of decisions has elaborated on appropriate survey technique.[17]

Novel Claims for New Technologies

When an advertiser makes a well-known claim, for example, "certified to reduce lead in water," there are well-established test standards that are widely relied on by reputable industry participants. However, in the case of product innovation or brand new technologies, no such standards are likely to exist.

Consider, for example, the iconic claim for the Dyson vacuum cleaner: "no loss of suction." At the time of the product claim debut, there were no published standards that directly

measured whether a vacuum cleaner did or did not lose suction over time. Dyson had to adapt existing standards from the IEC, among other entities, by taking a standard developed for instantaneous suction measurement and extending it to measure suction changes (if any) over time as the vacuum was gradually loaded with representative dust and dirt. This was not without controversy, as Dyson found itself embroiled in years of litigation with industry competitors who held differing views regarding how to measure this claimed attribute, and indeed differing views regarding what the claim even communicated to consumers.[18]

The Dyson litigation demonstrates the difficulty new technologies and novel claims can face, especially when the new technology competes with existing products.

The Problem of No Standards. There are different approaches to testing in support of such novel claims. When a company decides to communicate a new benefit or product attribute, its first stop in claims support is typically to scan the technical field for relevant, published standards by which it can test and verify the claim. As discussed in the Dyson example above, there may be relevant methods, but those methods may not be completely coextensive with the claim—creating the so-called problem of "lack of fit" between test and claim. The advertiser may, in such cases, seek to modify the standard to improve the fit. However, this can have several pitfalls—most importantly that any departure from the consensus standard can generate questions regarding the real-world relevance of the results. Departing from the standard necessarily also means that the test method no longer enjoys the consensus imprimatur of industry experts. In other words, when you depart from the standard, you are on your own to defend the departure.

Other advertisers may choose to develop entirely novel testing methods, either because no relevant standards exist or because they believe that a new test is simply better. For example, as mobile telecommunications technologies proliferated in the last decade, the competitors and outside vendors developed a variety of "drive testing" methods by which mobile network performance would be assessed on a nearly continuous basis as the tester systematically drove through neighborhoods.[19] The observed performance measurements were then "population-weighted" according to census data to arrive at combined performance metrics that could be compared from provider to provider. As those standards have evolved and been litigated, third-party vendors have developed and licensed testing services and claims, using proprietary variations on these methods. This has on occasion resulted in confusion as each such vendor has, occasionally, simultaneously endorsed claims of superiority for different industry competitors.

More recently, some telecommunications industry participants have gravitated to using crowdsourced data to support mobile download speed claims. These tests, such as those conducted by Ookla and OpenSignal, rely on consumers themselves to download a testing app to their mobile phones. Millions of ordinary consumers with those apps then ping known servers on the network to measure speeds, with the information feeding back to central databases for analysis. This kind of testing was endorsed for some purposes by NAD in the *T-Mobile USA* case, but later rejected when applied to the fixed broadband context just a year later in *Comcast Communications, Inc.*[20] Suffice it to say, the mobile and fixed broadband testing standards continue to evolve, with no broad consensus. At the same time, advertisers in these industries aggressively communicate speed superiority claims, arguing the merits of their respective test methods and data sets.

Substantiation of Novel Claims. It is a cliché that there is no "one size fits all" recipe for developing test methods that support novel claims. However, there are recognized hallmarks of good claims substantiation methodology that can improve the chances that the test will survive scrutiny in any FTC or NAD challenge. At the end of the day, a regulator or trier of fact will have to decide whether the test (1) meets the parameters of the claim being communicated, i.e., the test "fits" the claim; (2) has been performed in an unbiased fashion by people qualified to perform the test; (3) reflects what goes on in the "real-world" in the sense that it accurately predicts or models real world behavior; and (4) is reliable and repeatable, in the sense that its results can be replicated (within some predictable margin of error) if done again in the same lab or a different lab.

Identifying the Consumer Benefit. The starting point for substantiation is the identification of the consumer benefit being offered. With broader, multifaceted benefits, that analysis may not be straightforward. Consider, for example, the claim that a telecommunications network is the "best." There may be many different objective performance attributes relevant to overall network superiority such as download and upload speeds, latency, ability to maintain a connection without dropping, ability to connect on the first attempt, cost, coverage, and ease of access, among other factors. One could argue that in order to be "best," one would have to prevail on all or at least a majority of such factors as compared to the competition.[21]

What happens when the advertiser prevails on a majority of attributes but those "wins" are on factors that are less important to consumers than the attributes on which the advertiser lost? How does industry reach consensus as to what factors should matter? Should factors be weighted? The opportunity for mischief is great if the advertiser edits the list of relevant factors to the factors on which it prevails while conveniently omitting those attributes on which it loses.[22]

One way to arrive at an understanding of what consumers consider "best" is to conduct an attitudes survey, sometimes known as a Key Performance Indicator (KPI) survey. The survey asks current and prospective consumers of the product what factors are important to them in choosing products in the advertised category and may also ask consumers for a ranking or weighting of those factors (e.g., "Please indicate on

a scale of 1–7 how important [blank] is to you in deciding what [product] to purchase."). One can derive from the survey results a list of key attributes that drive purchase and their relative importance to one another. The list then can be related to objective testing results that are combined and weighted in accordance with the derived ranking.

One of the problems with relying on KPIs to derive a weighting of consumer relevance for product attributes is in the leading nature of the inquiry. If a survey instrument identifies attributes for ranking, it will inevitably include attributes that would not have been considered or mentioned by consumers had they not been specifically listed on the questionnaire. Moreover, if those attributes are disaggregated on an attribute-by-attribute basis in later product testing, they may not neatly be recombined and weighted for purposes of determining an "overall winner."

A preferred approach, although one that is not always feasible, is to provide the competing products to consumers and ask them to compare them under approximated real-world conditions and to assess which of the products they think is "better," either overall or with respect to specific attributes. This approach has the virtue of being direct, having real-world validity, and allowing consumers to decide for themselves, without artificial weighting or prompting, which product is "better" under any idiosyncratic definition of "better" they may hold. It also has the virtue of focusing only on consumer-relevant differences in the sense that small performance differences may either not be detected or may even out over the range of relevant attributes. The downside of this method, discussed below, is that such tests can obscure small but nevertheless significant and measurable performance improvements. In other words, the approach may not be sensitive enough to detect real differences, particularly as to novel benefits or those that are not consumer-perceptible, and may not be a preferred method for innovators seeking to tout product improvements.

Consumer Relevance. NAD has for years imposed a requirement that advertising claims be "consumer relevant" in the sense that the advertised benefit is meaningful to consumers. While NAD does permit an advertiser to tout even small improvements, it has made clear that such improvements must be relevant to consumers and that advertisers may not exaggerate their significance.[23]

The challenge for new and innovative products is establishing the consumer relevance of attributes or products that are previously unknown. When consumers are asked whether a product is "better" in some way, they must compare their experience with that product to one that is either tested side-by-side or otherwise saved in their memory. Consider a question such as, "Is this the best ice cream you have ever tasted?" Answering that question requires one to compare the ice cream in hand with a lifetime of memories of ice cream.

Difficulties arise, however, when the consumer has never before tasted ice cream. For new-to-the-world products, there may be no internal reference point to which a consumer could compare, and thus asking the consumer whether they think the product is the "best" has no meaning.

Referring back to the Dyson example discussed above, the "no loss of suction" attribute had never been advertised previously by any vacuum cleaner maker. Consumers were unfamiliar with the claim or its importance to product functioning. Dyson's earliest ads attempted to educate consumers about why having No Loss of Suction should matter when picking a new vacuum cleaner.

Dyson had enormous success in convincing people to understand the benefit by describing the frustrations that many consumers felt in having to buy new bags and then experience reduced suction over time as those bags filled with dust. Dyson's primary competitor at the time, which sold exclusively bagged vacuum cleaners and which enjoyed the revenue from selling bags, attacked the claim as effectively meaningless. What could "no loss of suction" possibly mean to consumers other than as being directly related to something with which they were already familiar: dust pickup? And, Hoover argued, its vacuums did better at dusk pickup from carpets than Dyson vacuums, despite not being able to claim "no loss of suction."[24]

In a similar fashion, in the early 2000s, Miller Brewing Company famously began advertising that Miller Lite had "More Taste" than Bud Light. The claim, which was supported by consumer testing of flavor intensity (not preference) was intended to communicate that Miller Lite had more flavor, which is important to consumers of light beers. Nevertheless, due to the novelty of the claim, Bud argued (ultimately unsuccessfully) that it would be interpreted by consumers as "better tasting."[25]

Testing Methods. A perennial testing issue is whether an attribute claim is best measured by "objective" testing as compared to consumer sensory testing. A full discussion of this issue is beyond the scope of this article. However, suffice it to say that each position can be argued for or against, on a highly fact-intensive basis.

In 1998, Clairol and L'Oreal engaged in a false advertising battle regarding claims of "gentleness" for hair coloring. Clairol's Hydience product claimed to be the "gentlest" hair coloring and the company substantiated that claim by virtue of two "objective" tests intended to measure the degree of damage caused by hair coloring to strands of hair—tensile strength testing, which measures the force required to stretched hair strands until breaking, and "combability," which measures the force required to drag a comb through treated hair. L'Oreal, by contrast, relied on both consumer and expert sensory measurements of hair quality and smoothness to rebut the claim. On a motion for preliminary injunction brought by L'Oreal, Clairol prevailed. The court found that Clairol's objective laboratory testing was sufficiently reliable and was not outweighed by the plaintiff's human sensory testing.[26]

However, before NAD in a recent case, a contrary view prevailed. There, the advertiser substantiated its claim to offer a tampon having "unsurpassed protection," supported

with a home use test in which women tried both competing products and then were asked which tampon, if either, protected them better. The challenger unsuccessfully argued that the claim was disproved by objective "leakage diary" tests in which women used both products and recorded on a daily basis any time they experienced a leak with a particular product. In the challenger's view, the product that leaked less frequently as objectively measured through the diary study could be said to provide better protection. In upholding the claim and "opinion" data on which it was based, NAD reasoned that "protection" was a multifactorial claim, and could depend on more than pure leakage. It concluded that women were capable of assessing protection reliably, without being told how to define the term.[27] On appeal, the National Advertising Review Board (NARB) refined NAD's decision by recommending that the advertiser indicate in the body of the claim whether it was based on consumer opinion.[28]

Testing Accuracy. Reliability, or agreement, deals with how close two measurements on the same subject are. Two measurements of the same subject may differ for a number of reasons, depending on the conditions under which the measurements were made. For example, a person stepping onto a bathroom scale two times in succession might obtain readings of 150 lbs. and 151 lbs. The differences, which most likely have to do with the measurement error inherent in the bathroom scale, can be characterized by means of common statistics.

Repeatability refers to the variation in repeat measurements made on the same subject under identical conditions. The measurements are made by the same instrument or method, the same observer (or rater), and typically over a short period of time, over which the underlying value can be considered to be constant. Observed variability in measurements made on the same subject in such a study can be attributed to measurement errors. This is a common metric of methodological accuracy, and has been incorporated into *Daubert*[29] jurisprudence.

Reproducibility refers to the variation in measurements made on a subject under changing conditions, such as conducting the measurement or experiment in different labs on different days. Changing conditions may be due to different measurement methods or instruments being used, measurements being made by different observers or raters, or measurements being made over a period of time, within which the "error-free" level of the variable could undergo non-negligible change.

However, these metrics are not the end of the story. A flawed test can generate the same flawed results time after time—i.e., "garbage in and garbage out." These hallmarks of tests function as checks on the test procedure in the sense that the test can be followed and replicated according to a recipe, and that test performance and measurements are not so variable as to cause unpredictable deviation in results. Some variance is always present, caused by factors such as lab instrument variability, operator variation, and slight differences in materials. What matters is that variation can be estimated with precision so that the experimenter can report results with an acceptable degree of confidence.

A test that is highly repeatable, reproducible, and reliable may also nevertheless be systematically biased against one product or another. Examples of arguments both for and against the presence of test bias abound in the cases.[30]

The Characteristics of Good Testing

In the Supreme Court's landmark *Kumho Tire* decision,[31] the Court noted that peer review is not a steadfast requirement of admissibility under *Daubert*: "It might not be surprising in a particular case, for example, that a claim made by a scientific witness has never been the subject of peer review, for the particular application at issue may never previously have interested any scientist."[32] However, it also cautioned that "on the other hand, [neither] does the presence of *Daubert*'s general acceptance factor help show that an expert's testimony is reliable where the discipline itself lacks reliability, as, for example, do theories grounded in any so-called generally accepted principles of astrology or necromancy."[33]

Daubert itself focuses appropriately on several factors of scientific reliability, such as:

- Whether a "theory or technique . . . can be (and has been) tested."
- Whether it "has been subjected to peer review and publication."
- Whether, in respect to a particular technique, there is a high "known or potential rate of error" and whether there are "standards controlling the technique's operation."
- Whether the theory or technique enjoys "'general acceptance'" within a "'relevant scientific community.'"[34]

For the purposes of ad claims substantiation for novel technologies, I would add at least the following:

- Follows industry standards to the extent possible where they exist, unless departures are technically justified.
- Testing procedures focus on consumer relevant attributes and the consumer benefit is clearly defined.
- Can be correlated to a reasonable degree with real-world results such that experimental data can be used to predict real world behavior within a reasonable margin of error.
- The tests are reliable, repeatable, and reproducible.
- Experts who are qualified in the fields of testing are willing to testify on behalf of the test. ■

[1] 81 F.T.C. 23 (1972). The FTC requires that advertisers have such substantiation in hand *before* the claims are disseminated in commerce, which is known as the "prior substantiation" rule.

[2] FTC Policy Statement Regarding Advertising Substantiation, *appended to Thompson Med. Co.*, 104 F.T.C. 648, 839 (1984), *aff'd*, 791 F.2d 189 (D.C. Cir. 1986).

[3] *See, e.g.*, Bureau of Consumer Protection, Fed. Trade Comm'n, Dietary Sup-

plements: An Advertising Guide for Industry 9 (1998) [hereinafter Guide], https://www.ftc.gov/system/files/documents/plain-language/bus09-dietary-supplements-advertising-guide-industry.pdf.

[4] Decision and Order, Mars Petcare US, Inc., FTC Dkt. No. C-4599 (Dec. 13, 2016), https://www.ftc.gov/system/files/documents/cases/161212_1523229_mars_petcare_us_decision_and_order.pdf.

[5] *See, e.g.*, Nice-Pak Prods., Inc., FTC Dkt. C-4556 (Nov. 2, 2015) (providing specific guidance regarding support for future claims of "flushability" on disposable wipes), https://www.ftc.gov/system/files/documents/cases/151102nice-pakdo.pdf.

[6] *See, e.g.*, *Mars Petcare*, *supra* note 4, at 2.

[7] *Pfizer, Inc.*, 81 F.T.C. at 64.

[8] *See, e.g.*, FTC v. Iovate Health Sci. USA, Inc., No. 10-cv-587 (W.D.N.Y. filed July 14, 2010); Beiersdorf, Inc., FTC File No. 092-3194 (filed June 29, 2011); Dannon Co., Inc., FTC File No. 0823158 (filed Dec.15, 2010).

[9] *See* J. Howard Beales, Timothy J. Muris & Robert Pitofsky, In Defense of the *Pfizer* Factors (Geo. Mason Univ. L. & Econ. Research Paper Series 12-49, May 2012), https://www.law.gmu.edu/assets/files/publications/working_papers/1249InDefenseofPfizer.pdf.

[10] *See* http://www.nsf.org/services/by-industry/water-wastewater/residential-water-treatment/residential-drinking-water-treatment-standards.

[11] *See* Zero Technologies LLC (ZeroWater Z-Pitcher), NAD Case Reports #5198 (July 2010).

[12] ASTM D5511-12, Standard Test Method for Determining Anaerobic Biodegradation of Plastic Materials Under High-Solids Anaerobic-Digestion Conditions (ASTM Int'l 2012); ASTM E1958-16A, Standard Guide for Sensory Claim Substantiation (ASTM Int'l 2016).

[13] Fed. Trade Comm'n, Statement of Basis and Purpose, Green Guides 75–112 (2012), https://www.ftc.gov/sites/default/files/attachments/press-releases/ftc-issues-revised-green-guides/greenguidesstatement.pdf. The lengthy discussion of the FTC's guidance on seals and certifications includes extensive citations to industry and other comments expressing concerns about the use of certifications from financially interested parties as a form of greenwashing. *See id.* at 77–88.

[14] *See, e.g.*, Press Release, Fed. Trade Comm'n, FTC Sends Warning Letters About Green Certification Seals (Sept. 14, 2015), https://www.ftc.gov/news-events/press-releases/2015/09/ftc-sends-warning-letters-about-green-certification-seals.

[15] *See, e.g.*, ASTM F608, Standard Test Method for Evaluation of Carpet Embedded Dirt Removal Effectiveness of Household/Commercial Vacuum Cleaners (ASTM Int'l 2017). The standard as published for many years included an appendix summarizing the results of field research testing that correlated lab to real-world results.

[16] 42 U.S.C. §1125(a).

[17] *See* Shari Seidman Diamond, *Reference Guide on Survey Research*, *in* Reference Manual on Scientific Evidence (Federal Judicial Center 3d ed. 2011) (collecting cases), https://www.fjc.gov/sites/default/files/materials/2017/SciMan3D01.pdf.

[18] *See, e.g.*, http://www.homeworldbusiness.com/dyson-sues-hoover-alleges-false-claims/.

[19] *See, e.g.*, T-Mobile USA, Inc. (More Data Capacity), NAD Case Report #5849 (May 2015).

[20] Comcast Commc'ns, Inc. (Xfinity Internet, Television and Telephone Services), NARB Panel 214 (Dec. 2016).

[21] *See, e.g.*, Procter & Gamble Co. (Pampers Baby-Dry Diapers), NAD Case Report #4599 (Nov. 2006) (parties agree that claim of superior overnight wetness protection supported by wins on two out of three objective tests of diaper performance).

[22] *See, e.g.*, Am. Home Prods. Corp. v. Johnson & Johnson, 654 F. Supp. 568 (S.D.N.Y. 1987) (enjoining such an edited ad because "where an advertisement is captioned 'Comparison of Non-Rx Analgesic Safety Profiles,' the reader might well be led to believe that the 'profile' has not been pruned to favor the advertiser, but is complete at least to the extent of including all of the side-effects as to which there is a significant difference between OTC analgesics").

[23] *See, e.g.*, Comcast Commc'ns, Inc. (Xfinity Internet, Television and Telephone Services), NAD Case Report #5974 (July 2016) (questioning relevance of fastest ISP rating, which is more reflective of consumer subscription preferences to certain Internet service tiers than to fastest speeds available).

[24] After years of litigation, Dyson and Hoover settled, with Dyson continuing to make its "no loss of suction" claim. The cyclone technology enabling this attribute has since become common in the industry.

[25] Miller Brewing Co. (Miller Lite and Miller Genuine Draft), NAD Case Report #4290 (Feb. 2005).

[26] Order, Clairol, Inc. v. L'Oreal, SA, No. 97-CV-00519 (D.D.C. Aug. 25, 1997).

[27] Kimberly-Clark Corp. (U by Kotex Sleek Tampons), NAD Case Report #5682 (Feb. 2014).

[28] Kimberly-Clark Corp. (U by Kotex Sleek Tampons), NARB Panel #194 (June 2014).

[29] Daubert v. Merrell Dow Pharm., 509 U.S. 579 (1993) (holding that the trial judge is the gatekeeper in assessing whether proffered expert testimony adheres to appropriate scientific methodology and should be admitted into evidence).

[30] *See, e.g.*, The Procter & Gamble Co. (Luvs with Nightlock), NAD Case Report #5767 (Sept. 2014) (challenger argument that diaper rewet test was systematically biased by test method).

[31] Kumho Tire Ltd. v. Carmichael, 526 U.S. 137 (1999).

[32] *Id.* at 151.

[33] *Id.*

[34] *Id.* at 149–50 (citing Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 592–94 (1993)).