Customer Confusion Audits:

Lessons from the Use of Consumer Confusion Evidence in the United States

Lauren E. Willis • Manoj Hastak • Jennifer King - 2018 (Updated September 2023)
CUSTOMER CONFUSION AUDITS:
LESSONS FROM THE USE OF CONSUMER CONFUSION EVIDENCE IN THE UNITED STATES

AUTHORS

KEY TERMS USED IN THIS REPORT

I. INTRODUCTION
   Motivation
   Confusion Audits and Caps

II. CONSUMER CONFUSION: LEGAL CONSTRUCT AND CAPS
   A. Consumer Confusion as a Legal Construct
      1. Creating Confusion or Profiting from Confusion
         • Broad Unfair Practices and Abusive Practices Prohibitions
         • Broad Deceptive Practices Prohibitions
         • Specific Prohibitions on Misleading Representations, including False Competitor Advertising
         • Informed Consent Requirements
         • Clear and Conspicuous Disclosure Requirements
      2. The Presence of Inaccurate Beliefs or the Absence of Knowledge
      3. The Persistence, Precision, and Depth of Consumer Knowledge
   B. Caps on Consumer Confusion
      1. Explicit Qualitative Caps and Implicit Quantitative Caps
      2. Barriers to Quantification
      3. Setting Customer Confusion Rate Caps

III. CONVENTIONAL METHODOLOGIES FOR ASSESSING CONSUMER CONFUSION IN U.S. LEGAL PROCEEDINGS
   1. Unaided Judicial Judgment
      • Construct Validity
      • Strengths
      • Weaknesses
         o Lack of Population Validity
         o Lack of Ecological Validity
         o Construct Validity Issues
         o Looming Impracticability
   2. Customer Witness Testimony
      • Strengths
         o Strong Ecological Validity
         o Can Produce Useful Qualitative Data
      • Weaknesses
         o Weak Population Validity
         o Potential for Measurement Error
      • The Use of Customer Witness Testimony Today
   3. Firm Behavior and Intent Evidence
      • Strengths
         o Strong Ecological Validity
         o Economical
      • Weaknesses
         o Looming Impracticability
4. Customer Behavior and Dissatisfaction Evidence

- **Strengths**
  - Strong Ecological Validity
  - Can Produce Useful Qualitative Data

- **Weaknesses**
  - Weak Population Validity
  - Potentially Ambiguous

5. A Note on Customer Satisfaction Evidence

IV. **SOCIAL-SCIENTIFIC METHODOLOGIES FOR ASSESSING CONSUMER CONFUSION IN U.S. LEGAL PROCEEDINGS**

A. The Use of Expert-Generated Evidence
   1. Admissibility
   2. Qualifications
   3. Ensuring Objectivity

B. The Four Main Types of Social-Scientific Evidence for Assessing Consumer Confusion
   1. Consumer and Customer Surveys (without a Stimulus)
      - **Construct Validity**
      - **Strengths**
        - Potential for Strong Population Validity
      - **Weaknesses**
        - Potential for Measurement Error
        - Potential for Data Analysis Errors
        - Logistical Challenges for Surveys of Actual Customers
   2. Survey and Usability Experiments (with a Control Condition)
      - **Construct Validity**
      - **Strengths**
        - Potential to Isolate Causation
        - Feasibility
      - **Weaknesses**
        - Looming Impracticability
        - Weak Ecological Validity and Looming Ecological Validity Challenges
        - Population Validity Concerns and Looming Challenges
        - Weak External Validity due to Subject Disinterest
        - Potential for Weak Internal Validity (Control Stimulus Construction Challenges)
        - Potential for Measurement Error
        - Potential for Data Analysis Errors
   3. Copy Tests and Usability Tests (without a Control Condition)
      - **Construct Validity**
      - **Strengths**
        - Feasibility
      - **Weaknesses**
        - Looming Impracticability
        - Weak Ecological Validity and Looming Ecological Validity Challenges
        - Population Validity Concerns and Looming Challenges
        - Weak External Validity due to Subject Disinterest
        - Potential for Measurement Error
        - Potential for Data Analysis Errors
   4. Expert Facial Analyses and Usability Inspections
• Construct Validity
• Strengths and Weaknesses
5. A Note on Methodological Imperfections and Convergent Validity

V. CONCLUSION
Lessons from the U.S. Experience
Looking to the Further Development of Customer Confusion Audits and Caps
AUTHORS

Lauren E. Willis is a law professor and William M. Rains Fellow at Loyola Law School, Los Angeles. Her research focuses on the intersection of consumer psychology, market structure, and law. She specializes in consumer financial regulation and is an expert on the use of disclosures, financial education, defaults, and performance-based regulation as tools for consumer protection. Prior to academia, she was a litigator handling consumer matters for the U.S. Federal Trade Commission and civil rights matters for the U.S. Department of Justice. She holds a J.D. from Stanford University.

Manoj Hastak is a marketing professor at American University’s Kogod School of Business. He has provided expertise to the Federal Trade Commission on deception and unfairness matters for over twenty years. In the course of that work, he has performed consumer surveys and experiments that assess consumer perception and confusion when consumers are faced with commercial marketing, advertising, and labeling, and with government-mandated disclosures. He has also served as an expert for the U.S. Food and Drug Administration, the U.S. Department of Justice, and the U.S. Consumer Financial Protection Bureau. He holds a Ph.D. in Marketing from the Pennsylvania State University.

Jennifer King is the Privacy and Data Policy Fellow at the Stanford Institute for Human-Centered Artificial Intelligence. She was previously the Director of Consumer Privacy at the Center for Internet & Society at Stanford Law School where she remains a non-resident fellow, and a product manager for a large technology company (Yahoo!). She has served as an expert on information architecture and the usability and user experience of websites and other electronic platforms, including for cases charging firms with unfair and deceptive trade practices. She holds a Ph.D. in Information Management and Systems from the University of California, Berkeley School of Information.
Key Terms Used in This Report

Key facts are facts about a transaction, the knowledge of which can potentially help consumers make decisions that will further their own welfare and will drive market competition over price and quality. These facts include whether the consumer has engaged in a transaction, the costs of the transaction, and limitations on the benefits the transaction provides.

Confusion audits are tests of a representative sample of a firm’s actual customers, performed or closely supervised by qualified independent experts, to ascertain whether the firm’s customers understand specified key facts about the transactions they have engaged in with the firm.

Confusion caps are limits on the proportion of a firm’s customers who demonstrate confusion about key facts relating to the transactions they have engaged in with the firm; exceeding these caps can demonstrate a violation of consumer protection or competition law.

Construct validity is the degree to which, apart from methodological issues (e.g., errors in data analysis), research identifies or measures what it purports or intends to identify or measure. For example, when the aim is to measure how well consumers understand key facts about a transaction, research that collects data about how well consumers think they understand that transaction may have weak construct validity because people are generally poor judges of the quality of their own knowledge and skills.

Population validity is the degree to which a sample used in research accurately represents the frequency and distribution of relevant traits of the population of interest as a whole, such that the results of the research using the sample reflect the results that would be obtained if the research were performed on the entire population of interest. For example, when the aim is to measure how well consumers understand key facts about a transaction in which they have engaged, research performed on a sample of individuals who have not engaged in that transaction has weak population validity.

Ecological validity is the degree to which the conditions under which research is performed accurately represent the relevant conditions under which the relevant population engages in the transaction of interest in the real world, such that the responses of study subjects are the same responses that occur in the real world. For example, when the aim is to measure how well consumers who engage in a transaction understand that transaction, research that asks subjects to read a disclosure and then tests them on their understanding of the transaction has weak ecological validity if in the real world, consumers engage in the transaction without reading the disclosure.

Firm behavior and intent evidence includes evidence about actions taken, knowledge possessed, and statements made by a firm or its management or employees during the ordinary course of the firm’s business. These types of evidence can be probative of a firm’s knowledge of customer confusion about key facts relating to transactions between the firm and its customers. Examples include market research performed by a firm to determine how consumers interpret marketing materials, user experience (“UX”) research performed by a firm to assess how consumers navigate the firm’s website, and emails among firm employees regarding customer confusion.

Customer behavior and dissatisfaction evidence includes evidence about actions taken and statements made by consumers. These types of evidence can be probative of consumer confusion about key facts relating to transactions engaged in by the consumer. Examples include: failure to use accounts opened or services purchased when the customer must use the account or service to benefit from it; purchase of insurance that replicates coverage the consumer already has; and customer complaints, requests for refunds, and chargebacks.
Unaided judicial judgment is the application of the court’s own judgment and common sense to ascertain how consumers in the real world would react to a stimulus, such as an advertisement. In U.S. law, this is often framed as an inquiry into the likely perceptions of “reasonable” or “average” consumers and is categorized as an “objective” approach.

An expert facial analysis is the application of an expert’s experience and knowledge of relevant principles established in the marketing and psychology literatures, to ascertain how consumers in the real world would react to and interpret a stimulus, such as an advertisement or disclosure. Among other things, the expert assesses whether consumers would be likely to be confused or misled by the stimulus. An expert facial analysis typically does not include testing with actual customers.

A usability inspection is the application of an expert’s experience and knowledge of relevant communication and user behavior principles established in the marketing, psychology, and human-computer interaction (HCI) literatures, to ascertain how consumers in the real world would react to and interact with an electronic interface, such as a website or mobile phone application. Among other things, the expert assesses whether consumers would be likely to be confused by the interface or to have difficulty using the interface to perform desired tasks. A usability inspection typically does not include testing with actual customers.

Copy tests involve showing research participants a stimulus, such as an advertisement, disclosure, or other communication from a firm to consumers, and asking the participants questions designed to elicit the messages and information the participants understand the materials to convey. Participants are generally asked about information and messages they derived from the stimulus during the test, not about their existing knowledge or beliefs about the product or service that is the subject of the stimulus communication.

Usability tests involve asking research participants to review the content of an electronic interface, such as a website or mobile phone application, and perform various tasks on that interface. A Human Computer Interaction expert observes how the participants attempt to use the interface and to what extent they are able to perform desired or assigned tasks. Participants are also asked questions about their subjective experience in using the interface and about the messages and information they understand the interface to convey. As with copy tests, participants are generally asked about their subjective experience using the interface during the test and about information and messages they derived from the interface during the test. Participants are generally asked neither about their prior experience with interfaces they may have encountered in the real world involving the same product, service, or firm, nor about their prior knowledge or beliefs about the product or service that the interface offers.

Survey experiments involve randomizing subjects into a treatment group and a control group and comparing the perceptions they draw from a test stimulus and a control stimulus through their answers to survey questions. The test stimulus could be, for example, an advertisement, disclosure, or other communication from a firm to consumers. The control stimulus is typically the same material shown to the treatment group, with the specific aspects of the communication that are suspected of being confusing or misleading corrected or eliminated. By observing any differences between the perceptions of the treatment group and the perceptions of the control group, the causal effect of the aspects of the firm-to-consumer communication that were suspected of being confusing or misleading can be assessed.

Usability experiments involve randomizing subjects into a treatment group and a control group and comparing the perceptions they draw from, and their ability to perform tasks on, a firm’s electronic interface (the test stimulus) and a control interface. Any differences between the perceptions and task performances of the two groups are attributed to the differences between the firm’s electronic interface and the control interface.

Consumer and customer surveys are used to ascertain the perceptions, knowledge, and beliefs about a transaction, product, or service held by consumers generally or by the customers of a particular firm. Rather than testing for perceptions, knowledge, and beliefs created by specified firm-to-consumer communications,
consumer and customer surveys test for perceptions, knowledge, and beliefs held by respondents based on their real world experiences, and, in the case of customer surveys, their actual interactions with the firm. In contrast to survey and usability experiments and tests, which aim to determine whether the particular materials tested confuse consumers, consumer and customer surveys aim to discover actual consumer or customer confusion, no matter the source of that confusion.
I. INTRODUCTION

The Australian Securities and Investment Commission (ASIC), Australia’s corporate, markets and financial services regulator, is assessing how to best further its mission to ensure fair, orderly and transparent consumer financial services and investment markets in light of twenty-first century trends in the marketplace. To this end, ASIC is considering the feasibility of confusion audits and caps as a form of performance-based regulation. This Report has been prepared to assist ASIC with this process. The Report reviews the U.S. experience with litigation and administrative proceedings1 aimed at deterring firms from creating or profiting from consumer confusion.2 More specifically, the Report:

- describes the consumer confusion constructs that are relevant to claims brought under U.S. consumer protection and fair competition law;
- identifies the rates of consumer confusion that have been found by U.S. courts to violate these laws;
- analyzes the strengths and weaknesses of the various types of evidence used to demonstrate actionable consumer confusion with respect to claims brought under these laws; and
- recounts the qualifications U.S. courts require of experts who perform social-scientific research used to demonstrate the presence or absence of actionable consumer confusion in legal proceedings.

Motivation

The animating concern behind the development of customer confusion audits and caps as a method of regulation is that consumers today are often confused about the costs and benefits of the market transactions in which they engage. They harbor affirmatively false beliefs about costs, they are unaware of limitations on benefits, and sometimes they do not realize they have engaged in a transaction at all. Modern consumer products and services, and financial products and services in particular, are bewilderingly complex, and the online channels through which they are increasingly sold only add to the confusion.

This confusion impairs consumer autonomy in the marketplace—the freedom to select transactions that satisfy personal needs within household financial constraints and to reject transactions that do not. It also dampens substantive competition—competition over true financial product and service quality and price—and fosters unfair competition—competition over which firm can best leverage consumer confusion to its advantage. This is not to say that eliminating confusion will lead to perfectly autonomous choices or a perfectly competitive marketplace, but it is one approach to increasing autonomy and substantive competition.

Regulators and enforcement agencies have had limited success in promoting consumer autonomy and substantive competition in today’s marketplace using traditional tools of consumer protection. Disclosure mandates have proven particularly ineffective. It is thus an opportune time to look to new regulatory methods that can keep pace with the modern marketplace.

A promising avenue in the consumer finance and retail investor space is to add performance-based regulation to the regulatory toolbox, as environmental, food safety, and other regulators have done. For

---

1 Legal proceedings in the U.S. include both administrative adjudications conducted within agencies, such as the Federal Trade Commission or the Consumer Financial Protection Bureau, and litigation conducted in the federal and state court systems. For ease of reference, this report will use the terms “litigation,” “court” and “judge” to cover judicial and administrative proceedings and adjudicators.

2 This Report uses “consumers” and “consumer products and services” to include retail investors and retail investment products and services.
example, in the U.S., the Food and Drug Administration requires pharmaceutical firms to show that actual customers understand the usage and dosing directions for a medication during a trial of over-the-counter sales before the firm can broadly sell the drug directly to consumers.\(^3\) Performance-based regulation identifies desired results and gives firms the necessary incentive and flexibility to achieve those results.\(^4\) One such desirable result would be to reduce consumer confusion about financial product and service costs and benefits. Recent work in economics has demonstrated that requiring firms to narrow the gap between consumers’ beliefs about financial products (specifically, consumer beliefs about key credit card and mortgage terms) and the true terms of those products, likely increases both consumer surplus and social welfare.\(^5\)

Some may question whether it is possible to narrow this gap, given the pervasiveness of consumer confusion, particularly regarding financial products and services.\(^6\) Reassuringly, mounting evidence demonstrates that modest changes in marketing, sales processes, and product design can diminish that confusion.

- In one study, providing price information in whole currency values rather than as a percentage rate helped low-financial-literacy consumers reduce the fees they paid for investments.\(^7\)
- Other research took a credit card sales website and added an interactive component—a slider that allowed consumers to immediately ascertain the amount of fees charged on transfers of balances of various sizes. Adding this slider increased the proportion of subjects who realized they would have to pay a fee to transfer a balance to the card from about a third to nearly three-fifths.\(^8\)
- In another experiment, eliminating vague, uninformative “low price” claims (“low APR,” “low fees,” and “low minimum payment”) from credit card marketing increased the rate at which the lowest cost, strictly dominating credit card was selected. Showing subjects a humorous video that contained actionable information (an explanation of where to locate interest rates and fees on a credit card disclosure) further increased the rate at which the best card was selected.\(^9\)

---


\(^4\) For an in-depth review of performance-based regulation, see Lauren E. Willis, Performance-Based Consumer Law, 82 University of Chicago Law Review 1309 (2015).


\(^6\) See, e.g., ASIC, Report 470, Buying Add-on Insurance in Car Yards: Why It Can Be Hard To Say No (2016), http://download.asic.gov.au/media/3549387/rep470-published-29-february-2016.pdf (finding massive confusion among Australian customers about what they purchased); Pete Lunn et al., PRICE Lab: An Investigation of Consumers’ Capabilities with Complex Products (May 2016), https://www.esri.ie/pubs/BKMNEXT306.pdf (finding consumer decision quality declines markedly if more than two or three attributes are considered); The PEW Charitable Trusts, Overdraft: Persistent Confusion and Concern About Bank Overdraft Practices 5 (June 2014), http://www.pewtrusts.org/-/media/assets/2014/06/26/safe_checking_overdraft_survey_report.pdf (finding more than half of U.S. bank account holders were unaware that they agreed to pay overdraft charges on debit and ATM transactions); AARP, 401(k) Participants’ Awareness and Understanding of Fees 6 (Feb. 2011), https://assets.aarp.org/rgcenter/econ/401k-fees-awareness-11.pdf (finding 71% of employer-sponsored defined contribution retirement plan participants in the U.S. falsely believe they pay no fees to their 401(k) provider to maintain their account).


\(^8\) Behavioral Insights Team, A behavioural approach to managing money: Ideas and Results from the Financial Capability Lab (May 2018).

\(^9\) Bruce I. Carlin et al., Millennial-Style Learning: Search Intensity, Decision Making, and Information Sharing, Management Science (2017), https://doi.org/10.1287/mnsc.2016.2689. Humor has been used effectively to increase attention and comprehension in other contexts as well. See, e.g., Dimuth Seneviratne & Brett R.C. Molesworth, Employing Humour and Celebrities to Manipulate Passengers’ Attention to Pre-flight Safety Briefing Videos in Commercial Aviation, 75 Safety Science 130 (2015) (finding experimentally that adding humor to airline safety video increased knowledge of the main safety rules. Further, people appear to be more likely to receive, remember, and use information given orally rather than in writing. See Niklos Gurney et al., Interview with the Vendor: Conversational
Pricing structure and product design can strongly influence how well consumers understand their transactions. Complex pricing schemes such as introductory “teaser” prices and multi-part “partitioned” prices routinely confuse consumers. Counterintuitive product designs can hamper consumer understanding. Again, various avenues have been shown to reduce consumer confusion. Examples include:

- **Simplifying prices** can reduce consumer confusion. For example, building a 15% fee into a product’s base price, displayed at the time consumers decide whether to place items into their online shopping carts, rather than adding the fee to the cart at the end of a website check-out process, apparently leads consumers to better appreciate the true cost of products. In one experiment, consumers spent 21% less on average when shopping on a website with the 15% fee built into the base price.\(^{11}\)

- **Eliminating counterintuitive features** from a product or service can also improve consumer understanding. When the U.S. Federal Reserve Board discovered that consumers did not understand credit card double-cycle billing (i.e., that a late payment on a credit card would trigger additional interest charges retroactively) or monthly payment allocation practices (i.e., that their credit card payments were being applied to their lowest interest rate balances first), the Board eliminated this confusion by banning these features.\(^{12}\)

This is not to say that regulation specifying the text of labels or marketing would be effective in decreasing consumer confusion. Time and again, mandated disclosures and warnings, and in particular disclosures for financial products, have been found ineffective.\(^{13}\) For example, every U.S. securities sales offering must disclose that “past performance is no guarantee of future results,” yet consumers stubbornly chase past returns in their investment decisions.\(^{14}\) Even disclosures that are somewhat effective in quiet laboratory settings are much less effective when more realistic environmental features are added. In one case, subject understanding of whether

---

Disclosures Improve Responsiveness to Strategically Withheld Information (July 2018, working paper, on file with author) (finding that listening to audio recording of oral disclosure rather than viewing a written, tabular-format disclosure of restaurant hygiene grades led to an over 25-percentage-point increase in accurate recall of the grades); Jeff Sovern, Written Notice of Cooling-Off Periods: A Forty-Year Natural Experiment in Illusory Consumer Protection and the Relative Effectiveness of Oral and Written Disclosures, 75 University of Pittsburgh Law Review 333 (2014) (based on surveys of firms, reporting that consumers are twice as likely to take advantage of three-day cooling-off periods to rescind a transaction when they are orally informed of the right to rescind, rather than only in writing (although rescission rates within the three-day period were very low for both groups)). In addition, subjects given information in writing in the form of a transcript of an oral conversation were more likely to recall and use overdraft fee information than subjects given the fee information in written tabular format. See Nikolos Gurney et al., Interview with the Vendor: Conversational Disclosures Increase Responsiveness to Strategically Withheld Information (July 2018, working paper, on file with author).

---


11. Tom Blake et al., Price Salience and Product Choice (April 2017). See also Raj Chetty et al., Salience and Taxation: Theory and Evidence, 99 American Economic Review 1145 (2009) (changing in-store price tags from base prices to tax-included prices dramatically affects consumer awareness of prices at the time they decide to buy the product; adding a 7.4% sales tax to price tags decreased demand by roughly the same amount (8%)).

12. See Ben S. Bernanke, Financial Innovation and Consumer Protection, Board of Governors of the Federal Reserve System (Apr. 17, 2009) (explaining that no matter how the Federal Reserve Board attempted to explain the credit card issuer practices of allocating payments to the portion of the balance carrying the lowest interest rate and of double-cycle billing, both were unintelligible to consumers, and therefore the Board banned these).


a home mortgage loan carried a fixed rate or an adjustable rate improved when a new mortgage disclosure form was reviewed in a quiet setting, but the addition of idle banter by the experimenter eliminated the improvement.15

Attempts to micromanage sales processes through regulation have similarly flopped. In the U.S., banks charge astronomical rates for overdraft services on checking accounts—over 7,000% in annualized terms—for what functions as a low-risk small loan.16 A default rule was established prohibiting banks from charging overdraft fees for debit and ATM transactions unless the accountholder opted into the bank’s overdraft program.17 The motivations for this change were research in other contexts demonstrating that default positions are sticky and a belief that consumers, particularly low-income consumers, were paying noncompetitive rates for unnecessary overdraft services.18 When the new rule was announced, many financial institutions engaged in marketing strategies targeted at new customers and at existing customers who had incurred overdraft fees in the past. These firms confused and harassed these customers into opting out of the new default in droves, leaving many low-income accountholders in no better position than before.19

Given the ease with which many marketing, sales processes, pricing structures, and product and service designs can be altered to reduce consumer confusion, it appears that many firms lack sufficient incentive to do so.20 Competitive forces can push firms in the opposite direction. Where it is financially advantageous to do so, a firm squeezed by competition and attending only to the bottom line is likely to confuse consumers and to profit from consumers’ confusion, sometimes unintentionally or even unknowingly.21

Well-enforced customer confusion caps—limits on the proportion of a firm’s customers who are confused about the costs and benefits of their transactions with the firm—would give firms an incentive to reduce customer confusion instead. In such a performance-based regulatory system, firms would be required to demonstrate that their customers know the key facts about the transactions they have engaged in with the firm. This demonstration could be made using confusion audits—tests of a representative sample of a firm’s actual customers—performed or supervised by qualified independent experts. Firms would be penalized (or rewarded) depending on the outcome of these audits.

To pass these audits, firms could increase the understanding of their customers through marketing campaigns and sales processes optimized for customer understanding. They could also simplify product and service prices and designs to comport with customers’ pre-existing beliefs about key product or service facts. Confusion caps would provide firms with the flexibility to choose among informational, pricing structure, and

16 See Lauren E. Willis, When Nudges Fail: Slippery Defaults, 80 University of Chicago Law Review 1155 (2013).
17 12 C.F.R. § 205.17.
18 74 Fed. Reg. at 59038-44.
20 Cf. Hooman Estelami, Consumer Perceptions of Financial Services Prices, in The Routledge Companion to Financial Services Marketing 239, 252 (Tina Harrison & Hooman Estelami, eds., Routledge: 2015) (detailing the lack of price transparency in the market for consumer financial services and suggesting, “the benchmarks by which management success is evaluated and rewarded may need to shift away from focusing only on measures of profitability and market share and include other measures that reflect the clarity and transparency of customer transactions”).
product simplification approaches to reducing their customers’ confusion. Firms could even micro-target tailored approaches to different types of customers and adapt their approaches over time as consumer knowledge shifts and product and service characteristics evolve.

<table>
<thead>
<tr>
<th><strong>KEY CONCEPTS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Facts:</strong></td>
</tr>
<tr>
<td><strong>Confusion Audits:</strong></td>
</tr>
<tr>
<td><strong>Confusion Caps:</strong></td>
</tr>
</tbody>
</table>

**Confusion Audits and Caps**

Confusion caps implicitly appear in U.S. law in a variety of contexts, including prohibitions on unfair, deceptive, or abusive practices, mandates requiring informed consent and clear and conspicuous disclosures, and prohibitions on unfair competition. An explicitly performance-based regulatory strategy for financial products aimed at reducing consumer confusion requires at least the following five steps:

1. Identification of key facts the knowledge of which would likely assist consumers in making autonomous decisions that drive markets toward substantive competition. The marketplace can be counted on to convey positive product and service attributes; confusion audits and caps are needed to improve understanding of prices and limitations on the benefits that products and services provide.

2. Clarity on the consumer confusion construct that is legally relevant. This would include: (a) whether the applicable law prohibits profiting from pre-existing confusion (mis-selling) or prohibits only identifiable actions by the firm that directly create new consumer confusion; (b) whether the law is aimed at lack of consumer knowledge or only at affirmatively inaccurate beliefs; and (c) the persistence (how strongly encoded in memory), precision, and depth of individual consumer knowledge with which the law is concerned.

3. Defining what proportion of a firm’s customers need to know the key facts about a transaction to facilitate consumer autonomy and motivate firms to substantively compete over that transaction feature.

4. Development of appropriate confusion audit methodologies for determining whether a firm’s customers understand or are confused about the key facts of their transactions with the firm.

5. Establishment of a system of rewards and/or penalties to incentivize firms to decrease their customers’ confusion about the key facts of their transactions with the firm.

This Report has been compiled to further inform ASIC about steps two through four. Specifically, this Report first sets forth the consumer confusion constructs at which U.S. consumer protection and fair competition laws are aimed. Next, the Report explains quantitative measures of consumer confusion that have been found in U.S. legal proceedings to violate the law, and discusses alternative approaches for establishing these customer confusion caps. Third, the Report reviews the strengths and weaknesses of various methodologies by which
U.S. courts ascertain consumer confusion, including both conventional and modern social-scientific approaches, and explains the qualifications experts need to produce admissible evidence using social-scientific methodologies. The Report briefly concludes with general lessons that can be drawn from the U.S. experience.

Throughout, the Report addresses the challenges posed by online consumer transactions. Firms are moving from mass marketing that employs a small set of marketing materials, standard purchase processes, gross pricing, and only a few product designs, to micro-targeted communications, electronic interfaces (e.g., websites, computer applications (apps), videos, and automated teller machine screens), pricing structures, and even product and service designs aimed at particular consumers in particular contexts. It is therefore becoming impracticable to analyze the effects on consumers of every firm communication or interaction. Indeed, micro-targeted machine-learning-generated marketing threatens to undermine the population and ecological validity of survey and usability experiments, copy and usability tests, expert facial analyses, and usability inspections. Instead, the focus must shift to the effects of this multitude of communications upon each firm’s actual customers.

* * * *

As firms increasingly harness big data and machine learning to personalize their interactions with consumers and to experiment with and adapt these interactions in real time, regulators and enforcement agencies need to develop new tools to police the marketplace. ASIC is uniquely positioned to develop and operationalize new approaches that can meet the demands of the modern marketplace for two reasons. First, ASIC is a leader in incorporating insights from behavioral economics into its work. It has conducted important research in this area, giving it key expertise in consumer financial decision-making and behavior. Second, ASIC is a savvy user of “regtech,” having already harnessed new technology to increase the speed and accuracy with which it can identify potential violations of consumer and investor protection regulations. Confusion audits and caps may be one new tool that has the potential to increase substantive competition in the consumer financial marketplace and to protect consumers.

---

22 For ease of reference, this Report uses “electronic interface” to include all such interfaces through which firms communicate to and interact with consumers, including those that may be invented in the future.

23 See, e.g., C. Perlich et al., Machine Learning for Targeted Display Advertising: Transfer Learning in Action, 95 Machine Learning 103, 104 (2014), doi 10.1007/s10994-013-5375-2 (“[M]achine learning ... play[s] a key role in the ad optimization process, particularly because of the simultaneous availability of (i) massive, very fine-grained data on consumer behavior, (ii) data on the brand-oriented actions of consumers, via instrumentation of purchase systems, and (iii) the ability to make advertising decisions and deliver advertisements in real time.... Each week, this learning system builds thousands of models to tally automatically, driving the advertising campaigns for major marketers across many industries.”); Ariel Ezrachi & Maurice E. Stucke, Artificial Intelligence & Collusion: When Computers Inhibit Competition, 2017 University of Illinois Law Review 1775, 1810 (2017) (reporting on the use by a large insurance company of an algorithm “which sought to optimize pricing by determining the likelihood that users would compare prices before purchasing insurance.”).


Limitations on this research:

In any research involving legal cases in the U.S., it must be understood that the results can be influenced by variation in public availability of data. Some cases are never reported, many unreported cases are not in accessible databases, and some case materials are filed under seal or subject to confidentiality agreements between the parties. For example, litigation brought by the Consumer Financial Protection Bureau has resulted in only a few expert reports pertaining to consumer understanding and confusion that are publicly available, and the Bureau has invoked exemptions for information relating to the supervision of financial institutions to deny requests under the Freedom of Information Act. The results here should be viewed with this caveat in mind.

In addition, this Report is not an instruction manual for how to perform research on consumer confusion. Handbooks designed to assist experts performing survey and human-computer interaction (HCI) research must be consulted, best practices must be adjusted to reflect the idiosyncrasies inherent in each research context, and tradeoffs must be made between achieving high levels of some types of methodological validity and high levels of other types of validity. No research has perfect internal validity and perfect external validity. Further, the fields of survey and HCI research are continually evolving.

Finally, the views herein are those of the authors and/or of the sources cited and should not be attributed to ASIC.

Style note: Throughout this Report, internal citations and quotation marks have been removed from quotations.
II. CONSUMER CONFUSION: LEGAL CONSTRUCT AND CAPS

A. Consumer Confusion as a Legal Construct

Consumer confusion is relevant to a number of federal and state legal claims in the U.S. The claims most pertinent to consumer protection and fair competition law are claims brought under statutes that: (a) broadly prohibit unfair and abusive practices; (b) broadly prohibit deceptive practices; (c) specifically proscribe false and misleading statements (including false competitor advertising); (d) require informed consumer consent; and (e) mandate clear and conspicuous disclosures.26 For all of these, the law primarily concerns itself with consumer confusion with respect to key facts. Three primary dimensions of consumer confusion as a measurable legal construct are: (1) whether the law aims at the creation of new confusion or at profiting from pre-existing confusion by firms; (2) whether the law aims at the presence of mistaken beliefs or the absence of knowledge on the part of consumers; and (3) the persistence, precision, and depth of consumer knowledge that the law seeks to produce.

1. Creating Confusion or Profiting from Confusion

| KEY CONCEPT: | Some U.S. consumer protection and fair competition laws prohibit creating consumer confusion (deception). Others prohibit profiting from consumer confusion, regardless of the source of that confusion (mis-selling). |

It is unlawful both for firms to create consumer confusion and for firms to take advantage of pre-existing consumer confusion. The line between creating consumer confusion and profiting from consumer confusion is admittedly hazy. When a firm generates confusion, it frequently also profits from that confusion. Further, a firm’s acts or omissions never confuse consumers in a vacuum; personal experience, cultural knowledge, and context all influence how consumers interpret the firm’s acts and omissions. It is thus unsurprising that the courts blur the concepts.

Yet, the distinction can be legally significant because creating confusion and profiting from confusion involve different causation issues and sometimes call for different forms of proof. When the legal question is whether a firm engaged in practices that led (or are likely to lead27) to consumer confusion, the court must determine whether the firm “caused” consumers to become confused. For these claims, the relevant construct is confusion among consumers that can be reasonably traced to specific practices of the firm.

When the legal question is whether a firm engaged in a practice that took advantage of consumer confusion, the court faces a different causation question: whether consumers’ confusion caused (or is likely to cause) consumers to take actions desired by the firm, such as buying a product, failing to cancel a service, incurring a fee, or the like. For these legal claims, the relevant construct is consumer confusion from which the firm profits, irrespective of the source of confusion.

The following discusses the primary types of legal claims that call for courts to examine consumer confusion, and whether creating or profiting from consumer confusion is the relevant construct involved in each type of claim.

---

26 This Report does not cover claims brought under trademark law, although much of the law regarding the admissibility of survey evidence comes from trademark litigation.

27 For claims brought by public enforcement authorities, the plaintiff typically does not need to demonstrate actual consumer confusion but can instead prevail by establishing that confusion is likely.
• **Broad Unfair Practices and Abusive Practices**\(^{28}\) Prohibitions

The U.S. Federal Trade Commission Act prohibits “unfair ... acts or practices in or affecting commerce.”\(^{29}\) The Consumer Credit Protection Act expands this prohibition to include “abusive acts or practices” in connection with consumer financial products and services.\(^{30}\) Only government enforcement authorities can bring claims for violations of these federal statutes. The securities laws similarly prohibit investment advisers from engaging in “any act, practice, or course of business which is fraudulent ... or manipulative.”\(^{31}\) Most states have statutes containing similar prohibitions; most of these permit private enforcement by consumers, in addition to public enforcement.\(^{32}\) The relevant confusion construct for both unfair practices claims and for abusive practices claims is confusion from which the firm profits, regardless of the source of that confusion.

A business practice is **unfair** under U.S. federal law if it “causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition.”\(^{33}\) “Substantial injury” includes small harms inflicted on a large number of people.\(^{34}\) “Reasonably avoidable” means that “consumers have reason to anticipate the impending harm and the means to avoid it....”\(^{35}\) Unfairness actions are brought “to halt some form of seller behavior that unreasonably creates or takes advantage of an obstacle to the free exercise of consumer decision making.”\(^{36}\) The Federal Trade Commission has identified several prototypical examples of such behavior, including firms that “withhold or fail to generate critical price or performance data ... leaving buyers with insufficient information for informed comparisons.”\(^{37}\)

Confusion over the key facts regarding a market transaction is an obstacle to the free exercise of consumer decision making. Therefore, both producing confusion as well as taking advantage of confusion are unfair practices, if consumers are thereby injured. Confused consumers have no reason to anticipate impending harm and therefore are not equipped with the means to avoid it. Resultant injuries to consumers will rarely be outweighed by benefits to consumers, given that without consumer understanding, substantive competition over the misunderstood cost or benefit will not take place.

---

\(^{28}\) Some statutes use the terms “manipulative” or “unconscionable” practices instead of abusive. For ease of reference, this Report uses “abusive practices” to encompass all of these.


\(^{30}\) 12 U.S.C. § 5531 (a part of the “Dodd-Frank” legislation passed by the U.S. Congress in the wake of the 2008 financial crisis).


\(^{34}\) See, e.g., Federal Trade Commission v. Commerce Planet, Inc., 878 F. Supp. 2d 1048, 1078 (C.D. Cal. 2012), affirmed in relevant part, vacated in part, remanded, 815 F.3d 593 (9th Cir. 2016), and affirmed in part, 642 F. App’x 680 (9th Cir. 2016) (holding that a small fee, when assessed on many consumers, can constitute “substantial injury”).

\(^{35}\) Consumer Financial Protection Bureau v. Navient Corp., No. 3:17-CV-101, 2017 WL 3380530, at *21 (M.D. Pa. 2017). Further, a court will not assume that consumers “understood the disclosures that were made to them so that they had ‘reason to anticipate the impending harm and the means to avoid it.’” Id.


\(^{37}\) Id.
Many of the practices engaged in by financial product and services providers that have been deemed unfair have involved profiting from consumer confusion:

❖ In one case, the Consumer Financial Protection Bureau charged a firm with unfair practices for having imposed a recurring fee that was not clearly explained to its customers, such that a substantial proportion of its customers did not know it would be imposed.  

❖ Several cases have held that lenders or affiliates of lenders that trick or deceive consumers into borrowing have thereby committed an unfair practice. In one of these cases, a mortgage refinance service committed an unfair practice when it signed up customers by misrepresenting the success of its loan program in preventing loan default or foreclosure and misrepresenting various aspects of the program’s price. The court held that creating the customer confusion was deceptive, and taking advantage of that confusion was unfair.

❖ Another case held that directing students to expensive loans would constitute an unfair practice when “some students ... did not even realize they took out the [loans] because of the rushed and automated manner in which [the defendant, a for-profit school] processed ... students’ paperwork.”

Because profiting off of confusion is unfair, the relevant consumer confusion includes any confusion regarding the key facts of the transaction, irrespective of the source of confusion.

An abusive practice includes a practice that “takes unreasonable advantage of (A) a lack of understanding on the part of the consumer of the material risks, costs, or conditions of the product or service [or] (B) the inability of the consumer to protect the interests of the consumer in selecting or using [the product or service].” Courts have held that “to take advantage of” means “to make use of for one’s own benefit” or “to profit by.” Additionally, courts have held that inability to protect a consumer’s own interests includes “oppressive circumstances, such as when a consumer is unable to protect herself not in absolute terms but relative to the excessively stronger position of the defendant.”

Consumers who do not understand material risks, costs, or conditions of a product or service generally cannot protect their own interests. Selling a financial product to these confused consumers takes advantage of both their inability to protect their interests and their underlying lack of understanding.

❖ In one case, a used car auto dealer and financer lured consumers with misleading advertising about loan costs, a practice found to be deceptive. The dealer/financer then kept consumers confused about the true cost of financing the cars they were buying so as to sell the car loans to the consumers. The dealer/financer thus profited from consumers’ confusion and resultant inability to protect their

---

40 Consumer Financial Protection Bureau v. ITT Educational Services, Inc., 219 F. Supp. 3d 878, 915-16 (S.D. Ind. 2015); see also Consent Order, In re Amerisave Mortgage Corp., File No. 2014-CFPB-0010, 08/12/2014, ¶¶ 64–89 (explaining that fostering mistaken borrower beliefs that a third-party service provider was independent, when the provider was an affiliate of the lender, was deceptive, and taking advantage of that confusion by allowing the affiliate to charge as much as a 900% markup on its services, was unfair).
41 12 U.S.C. § 5531(d)(2). Some state consumer protection statutes retain an old common law intent requirement, defining taking advantage of consumer ignorance as an unconscionable practice only if the business knew or should have known the consumer’s ignorance. See, e.g., Ohio Rev. Code § 1345.03 (listing as factor to consider is determining whether act is unconscionable, in violation of state UDAAP law, “[w]hether the supplier has knowingly taken advantage of the inability of the consumer reasonably to protect the consumer’s interests because of the consumer’s ... ignorance”). Today, given the amount of consumer testing that firms perform, it is increasingly likely that firms will know when their customers are ignorant or confused about a key fact pertaining to the product or service.
interests, a practice found to be abusive.\textsuperscript{43} Although in this example, the firm created the confusion that it exploited, because the prohibited practice is taking advantage of consumers’ inability to protect their own interests, the relevant confusion could be from any source.

- \textit{Broad Deceptive Practices Prohibitions}

In addition to unfair and abusive practices, the U.S. Federal Trade Commission Act and the Consumer Credit Protection Act prohibit deceptive acts and practices.\textsuperscript{44} Again, only government enforcement authorities can bring claims for violations of these federal statutes. The securities laws similarly prohibit investment advisers from engaging in deceptive acts and practices.\textsuperscript{45} Most states have statutes containing similar prohibitions; most of these permit private enforcement by consumers, in addition to public enforcement.\textsuperscript{46}

A \textit{deceptive} practice is “a representation, omission or practice that is likely to mislead the consumer acting reasonably [under] the circumstances, to the consumer’s detriment.”\textsuperscript{47} Because it is the effect on the consumer that matters, a literally true statement can be deceptive,\textsuperscript{48} and a literally false statement can be non-deceptive.\textsuperscript{49} An omission, such as the failure to convey a limit on an advertised benefit, can sometimes be

\textsuperscript{43} Consent Order, In the Matter of Y Kings Corp., also doing business as Herbies Auto Sales, File No. 2016-CFPB-0001, 01/21/2016, ¶¶ 70-76.

\textsuperscript{44} 15 U.S.C. § 45 (a); 12 U.S.C. § 5531.


\textsuperscript{47} See Federal Trade Commission Policy Statement on Deception, 103 F.T.C. 174, 175 (1984). Under the Fair Debt Collection Practices Act, the standard is whether the act or omission would mislead an “unsophisticated” consumer rather than a “reasonable” consumer. See, e.g., McMahon v. LVNV Funding, L.L.C., 744 F.3d 1010 (7th Cir. 2014); Clomon v. Jackson, 988 F.2d 1314 (2d Cir. 1993) (applying “least sophisticated consumer” standard to FDCPA cases).

\textsuperscript{48} See, e.g., State ex rel. Miller v. Rahmani, 472 N.W.2d 254, 258 (Iowa 1993) (“A statement which is literally true may nevertheless be found to be deceptive.”); Donaldson v. Read Magazine, Inc., 333 U.S. 178, 188 (1948) (“Advertisements as a whole may be completely misleading although every sentence separately considered is literally true. This may be because things are omitted that should be said, or because advertisements are composed ... in such way as to mislead.”). For a typology that categorizes true but misleading claims on the basis of the underlying psychological mechanisms driving consumers to misinterpret these claims, see Manoj Hastak & Michael B. Mazis, Deception by Implication: A Typology of Misleading Advertising and Labeling Claims, 30 Journal of Public Policy & Marketing 157 (2011), https://doi.org/10.1509/jppm.30.2.157.

\textsuperscript{49} See, e.g., Muha v. Encore Receivable Management, Inc., 558 F.3d 623, 627 (7th Cir. 2009) (“If a statement would not mislead the unsophisticated consumer, it does not violate the [Fair Debt Collection Practices Act], even if it is false in some technical sense. For purposes of [the statute’s prohibition on false statements], then, a statement is not ‘false’ unless it would confuse the unsophisticated consumer.”).
deceptive.\textsuperscript{50} Small print disclaimers generally do not cure deception; given how rarely consumers read these, it would be unreasonable to expect consumers to do so.\textsuperscript{51}

One might characterize deceptive acts as involving confusion created by an act or practice of the firm and deceptive omissions as involving pre-existing confusion that the firm does not correct and from which the firm profits. Yet the distinction is not entirely clean in practice. In assessing whether an act such as making a representation is likely to mislead consumers, U.S. courts assess the overall net impression that reasonable consumers are left with when they interact with the firm.\textsuperscript{52} The relevant source of consumer confusion is thus an amalgam of pre-existing consumer beliefs and an act or representation of the firm. Omissions are deceptive when consumers’ pre-existing beliefs lead them to interpret the firm’s silence (or failure to effectively communicate, such as by the use of fine print) as communicating information, although one might characterize the defendant’s silence as an act.

For example, an “act” might be sending a consumer debtor a letter suggesting a settlement of a debt. An “omission” might be failing to further convey in that letter that the debt collector cannot legally collect on the debt because it is time-barred. It is not entirely comprehensible to say it is the act and not the omission or the omission and not the act that causes a consumer reading the letter to incorrectly believe that the debt is legally collectible. The act of sending the letter impliedly represents that the debt is collectible only because consumers already believe that debt collectors do not offer to settle uncollectible debt, even though the letter is silent on this point. The failure to convey that the firm cannot collect the debt is a deceptive omission only because consumers already incorrectly believe that debt collectors do not offer to settle uncollectible debt.

- In one case brought by the Consumer Financial Protection Bureau, the court found that a debt collector engaged in deceptive practices by “servicing and collecting on [unenforceable loans]” because doing so “created the ‘net impression’ that the loans were enforceable and that borrowers were obligated to repay the loans in accordance with the terms of their loan agreements” even though those “agreements” were void.\textsuperscript{53}

- A case pursued by the Federal Trade Commission involved a similar type of omission. The defendant firms capitalized on the widespread consumer belief that a “credit card” is a general purpose card unless it is specifically identified as a store card that can be used only for purchases at a particular merchant. The appellate court reviewing the case held that by representing that the card had the

\textsuperscript{50} In re National Credit Management Group, L.L.C., 21 F. Supp. 2d 424, 441 (D.N.J. 1998) (“[T]he omission of material information, even if an advertisement does not include falsehoods, may result in a violation of [the Federal Trade Commission Act].”); Sterling Drug, Inc. v. Federal Trade Commission, 741 F.2d 1146, 1154 (9th Cir. 1984) (holding that an omission can be deceptive); Bain v. Metropolitan Mortgage Group, Inc., 175 Wash. 2d 83, 116 (2012) (same).

\textsuperscript{51} See Yannis Bakos et al., Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts, 43 Journal of Legal Studies 1 (2014), https://doi.org/10.1086/674424 (finding that fewer than two in one thousand consumers click on fine print terms in online software contracts, and those who do spend too little time, on average, to read more than a small portion of the text). See also, e.g., Federal Trade Commission v. Johnson, 96 F. Supp. 3d 1110, 1146 (D. Nev. 2015) (“The mere fact that the sites contained disclosures in smaller print and described the upsells as “bonuses” and trials at the bottom of the order pages, does not alter the deceptive net impression as to the cost and nature of the product because consumers would not be inclined to seek out this information.”; citing expert testimony that consumers ignore what they perceive to be “legal text”); Removatron International Corp. v. Federal Trade Commission, 884 F.2d 1489, 1497 (1st Cir. 1989) (“Disclaimers or qualifications in any particular ad are not adequate to avoid liability unless they are sufficiently prominent and unambiguous to change the apparent meaning of the claims and to leave an accurate impression. Anything less is only likely to cause confusion by creating contradictory double meanings.”).

\textsuperscript{52} See, e.g., Federal Trade Commission v. Cyberspace.Com, L.L.C., 453 F.3d 1196, 1200 (9th Cir. 2006) (“A solicitation may be likely to mislead by virtue of the net impression it creates even though the solicitation also contains truthful disclosures.”); Federal Trade Commission v. Washington Data Resources, 856 F. Supp. 2d 1247, 1272 (M.D. Fla. 2012), affirmed, 704 F.3d 1323 (11th Cir. 2013) (“To be considered reasonable, the interpretation or reaction does not have to be the only one. When a seller’s representation conveys more than one meaning to reasonable consumers, one of which is false, the seller is liable for the misleading interpretation.”).

features of a general purpose card, the defendants created “the deceptive impression that they were offering consumers a general purpose credit card.” In affirming the trial court’s finding of deception, the appellate court noted that the defendants “failed to dispel the confusion that [their] representations created among reasonable consumers.” \(^{54}\)

Another example would be clearly disclosing some but not all costs of a financial product. This is a deceptive practice because consumers are likely to assume, based on pre-existing experience, that all costs will be clearly disclosed together. \(^{55}\) Quite a number of cases involve deception carried out in this fashion:

- A failure to prominently disclose a 1% currency conversion fee on monthly credit card statements was held to be a deceptive practice. Although the fee was disclosed in the small print cardholder agreement, all other charges were disclosed on the monthly statement, so consumers would expect the charge to be disclosed on the statement too. \(^{56}\)

- In a case finding a payday lender to have engaged in deceptive practices, the court explained that the loan disclosures gave “the impression that a $300.00 loan from the Lending Defendants will only cost borrowers $90.00.” However, due to the automatically renewing structure of the lender’s loans, “unless borrowers read the fine print and take the necessary steps to opt out of the renewal plan, such a loan will incur $675.00 in fees.”\(^{57}\)

- In another case, a firm sent consumers unsolicited checks with the condition attached, but not clearly disclosed, that cashing the check would enroll consumers in a subscription service. In finding defendants to have engaged in a deceptive practice, the court explained: “[T]he receipt of a check, the perusal of which would reveal no obvious mention of an offer for services, no product information, and no indication that a contract is in the offing, coupled with an invoice that has no advertising or solicitation purpose, creates an overall impression that the check resolves some small, outstanding debt.” \(^{58}\)

As these examples demonstrate, the relevant consumer confusion construct in deception claims encompasses both confusion stemming from acts of the firm and pre-existing confusion from any source when the firm (by omission) fails to take action that would dispel the confusion.

**Specific Prohibitions on Misleading Representations, including False Competitor Advertising**

In addition to general prohibitions on unfair, deceptive, or abusive practices, other statutes in the U.S. specifically prohibit *false or misleading representations* in various contexts. Depending on the statute, violations of these specific prohibitions can be pursued by (a) public enforcement agencies, (b) consumers, or (c) one competitor against another. Misleading consumers can also be (d) an element of another claim, such as an antitrust claim.

Examples of *misleading representation prohibitions enforceable by (a) agencies or (b) consumers* abound, particularly in the consumer finance arena. For example, federal and state statutes specifically prohibit false and misleading statements in advertising, including advertising related to consumer financial transactions. \(^{59}\)

---

\(^{54}\) Federal Trade Commission v. USA Financial, L.L.C., 415 F. App’x 970, 973 (11th Cir. 2011).


\(^{59}\) See, e.g., 15 U.S.C. § 52 (specifically banning false advertisements and declaring dissemination of these to be unfair or deceptive practice; defining false advertisement to include false and misleading statements and omissions); Cal. Bus. & Prof. Code § 17200.
The Fair Debt Collection Practices Act prohibits the use of “any false, deceptive, or misleading representation or means in connection with the collection of any debt,” including, e.g., “threat[ening] to take any action that cannot legally be taken.” The Credit Repair Organizations Act prohibits credit repair organizations from making or using any untrue or misleading representations about the services of the credit repair organization. Another example comes from laws protecting investors; regulations under the Investment Advisers’ Act prohibit false and misleading advertising by registered investment advisers.

Most courts treat the question of whether a representation is misleading in the context of prohibitions on specific misleading representations that are enforceable by consumers as presenting a question of law. The underlying facts—the text and context of the representation—are questions of fact, but whether those facts make the representation misleading is, in most U.S. jurisdictions, a question of law decided by a court not a jury. Further, these jurisdictions do not always require extrinsic evidence (e.g., a consumer survey or expert facial analysis) to make that determination and instead often rely on unaided judicial judgment.

A few courts hold that whether a communication from a debt collector is misleading under the Fair Debt Collection Practices Act is a question of fact. Unless the dunning letter is, in the court’s subjective opinion, so plainly nonmisleading or plainly misleading that no reasonable jury could find otherwise, the plaintiff must submit expert-generated survey evidence to prove that the letter violates the Fair Debt Collection Practices Act. Courts have stated that they treat prohibitions on debt collector misrepresentations differently from other types of misrepresentation claims because for debt collection matters the applicable standard is whether an “unsophisticated consumer” or the “least sophisticated consumer” would likely be misled, rather than the usual “reasonable consumer” standard. One court has explained:

How a particular notice affects its audience is a question of fact, which may be explored by testimony and devices such as consumer surveys.... District judges are not good proxies for the “unsophisticated consumers” whose interests the statute protects. Unsophisticated readers may require more explanation than do federal judges; what seems pellucid to a judge, a legally sophisticated reader, maybe opaque to someone whose formal education ended after sixth grade. To learn how an unsophisticated reader reacts to a letter, the judge may need to receive evidence.

Of course, what seems pellucid to a judge might also be opaque to even an average consumer who is not particularly unsophisticated. However, there are practical reasons not to require individual consumers to
present expert survey evidence—the amount that can be recovered by any one consumer is usually insufficient to justify the expert witness expense.

Regardless of whether the issue is one of law or fact, the consumer confusion construct at issue in these cases is confusion caused by the firm’s misrepresentations. Taking advantage of consumer confusion is prohibited by other provisions of most of the same statutes that specifically prohibit misrepresentations, but the misrepresentation prohibitions focus on whether consumers are likely to be misled by statements of a firm.

In category (c) are fair competition statutes that permit firms to pursue claims against one another for misleading promotional materials. These laws are intended to protect firms from false competitor advertising that unfairly draws customers to the competitor and away from the plaintiff. A statement in commercial advertising or product labeling is misleading if it conveys a false impression, is misleading in context, or is likely to deceive consumers. The misrepresentation is actionable if it is about the firm’s own product or is about a competitor’s product. For an omission to form the basis of a false competitor advertising claim, the omission must be “relevant to an affirmative statement that is made false or misleading by its omission.” As in the context of deceptive practices claims, a literally false statement can be non-misleading, and a true statement can be misleading.

To bring a false competitor advertising claim, a firm must show that its competitor’s activities deceived consumers. The relevant legal construct is consumer confusion created by the competitor’s marketing or labeling, and not confusion from another source. Presumably any background confusion benefits all competitors equally. For example, consumers may incorrectly believe that prices for credit are based on creditworthiness and not on the type of loan or the amount of credit. An omission to form the basis of a false competitor advertising claim, the omission must be “relevant to an affirmative statement that is made false or misleading by its omission.”

Whether a firm has misled consumers can also be (d) an element of other legal claims. For example, in one antitrust case, the issue posed was whether a telephone services provider had violated antitrust laws by leading its customers to mistakenly believe that they could obtain a particular type of ancillary service only from it. If the provider had so misled consumers, competitors seeking to sell the same service would be prevented from doing so, because disabusing consumers of this belief was a public good and no one competitor that engaged in a corrective advertising campaign could be sure it would reap the business produced from consumers who

---

67 See, e.g., Johnson v. Revenue Management Corp., 169 F.3d 1057, 1063 (7th Cir. 1999) (Eschbach, J., concurring) ("[A] system which places this additional cost [of conducting surveys] on litigants will make the cost of filing suit under the [Fair Debt Collection Practices Act] prohibitive.").

68 The relevant federal statute here is the Lanham Act, 15 U.S.C. § 1051 et seq., and states often have similar statutes. The Lanham Act also prohibits trademark infringement and dilution.


71 See, e.g., Schering-Plough Healthcare Prod., Inc. v. Schwarz Pharmaceuticals, Inc., 586 F.3d 500, 512 (7th Cir. 2009) ("The purpose of the false-advertising provisions of the Lanham Act is to protect sellers from having their customers lured away from them by deceptive ads . . . or labels, or other promotional materials. . . . Many literally false statements are not [misleading]."); Abbott Laboratories v. Mead Johnson & Co., 971 F.2d 6, 15–16 (7th Cir. 1992) ("[O]ne business's] statement that [its product] has a lower osmolality than [the product of its competitor], while literally true, is misleading because the difference in osmolality has no therapeutic significance").

72 See, e.g., Johnson & Johnson-Merck Consumer Pharmaceuticals Co v Rhone-Poulenc Rorer Pharmaceuticals, Inc., 19 F.3d 125, 129–30 (3rd Cir 1994) (explaining that, unlike the Federal Trade Commission when it brings suit as a plaintiff, a Lanham Act plaintiff "cannot obtain relief by arguing how consumers could react; it must show how consumers actually do react") (emphasis in original).

were no longer confused. The effect would be to leave the current provider in a monopoly position, from which it could charge inflated monopoly prices. The relevant legal construct in these types of cases is confusion among the firm’s customers that was induced by specific statements or acts of the firm.

- **Informed Consent Requirements**

A variety of U.S. laws, including consumer financial protection laws, require firms to obtain informed consumer consent. If consumers are confused, any consent the consumer gives is uninformed. Thus, the firm that accepts such consent is in violation of the legal mandate to obtain informed consent.

Whether a consumer gave informed consent to a particular transaction cannot always be determined by analyzing particular materials given by the firm to the consumer. Consumers may not have seen, read, or processed materials produced by the firm. Consumers can become informed through many sources, and informed consent laws do not require that consumers become informed through a particular source. On the other hand, where the firm is likely to be the dominant source of information, judges have found it reasonable to infer that if the firm did not disclose the information, the firm’s customers are unlikely to have been informed.

That a firm did disclose the information in some fashion, however, is not sufficient to demonstrate informed consumer consent. For example, investment advisers in the U.S. are often permitted to provide investment advice to retail clients in the face of a conflict of interest, so long as the adviser has obtained informed client consent to the conflict. Client consent can only be informed if the client understands the existence and nature of the conflict of interest, and so the conflicted adviser must dispel any client confusion on both points. The Securities and Exchange Commission has made clear that this informed consent requirement is not merely a disclosure requirement—it is an affirmative requirement to ensure clients understand the conflict:

> [N]o hard and fast rule can be set down as to an appropriate method for [the adviser] to disclose the fact that she proposes to deal on her own account. The method and extent of disclosure depends upon the particular client involved. The investor who is not familiar with the practices of the securities firm requires a more extensive explanation than the informed investor. The explanation must be such, however, that the particular client is clearly advised and understands before the completion of each transaction that [the adviser] proposes to sell her own securities.


75 This is in contrast to clear and conspicuous disclosure requirements, which can be violated even if the consumer does not need the disclosure because the consumer is already well-aware of the facts being disclosed. Cf. Washington v. Internet Order, LLC, No. C14-1451JLR, 2015 WL 918694, at *5 (W.D. Wash. Mar. 2, 2015) (distinguishing between consumer protection claims for a firm’s misrepresentations or failure to provide consumers with clear and conspicuous disclosures from claims against that firm for failure to obtain informed consumer consent to a transaction).

76 See, e.g., Federal Trade Commission v. Credit Bureau Center, LLC, No. 17 C 194, 2018 WL 3122179, at *7 (N.D. Ill. June 26, 2018) ("A website that fails to provide a consumer any information about a service cannot obtain a consumer’s express informed consent to purchase that service."); Federal Trade Commission v. DirecTV, Inc., No. 15-CV-01129-HSG, 2016 WL 5339797, at *3 (N.D. Cal. Sept. 23, 2016) (examining advertisements rather than evidence about consumer knowledge or confusion to assess whether "consumers did not have sufficient information and thus could not have given informed consent").

77 See, e.g., In the Matter of Arleen W. Hughes, 27 S.E.C. 629 (1948), affirmed, Hughes v. Securities and Exchange Commission, 174 F.2d 969 (D.C. Cir. 1949) ("Registrant has an affirmative obligation to disclose all material facts to her clients in a manner which is clear enough so that a client is fully apprised of the facts and is in a position to give his informed consent.... Registrant cannot satisfy this duty by executing an agreement with her clients which the record shows some clients do not understand.").
The relevant legal construct is whether clients are confused about the conflict at the time they engage in transactions in which the adviser has a conflict of interest; the source of that confusion is inconsequential.

Another context in which a firm must obtain informed consumer consent is when the firm is charging consumers pursuant to a negative option plan.78 The Telemarketing Sales Rule requires firms in telephone transactions to obtain “the express informed consent of the customer” to a negative option feature before billing the customer.79 Similarly, the Restore Online Shoppers’ Confidence Act requires firms in online transactions to “obtain[] a consumer’s express informed consent before charging the consumer’s credit card, debit card, bank account, or other financial account for products or services [through a negative option plan].”80 Firms have the responsibility to ensure that consumers understand they are enrolling in a negative option membership plan.81 If reasonable consumers do not understand this, regardless of the source of the confusion, the firm has failed to obtain the required informed consent.

Therefore, for laws mandating informed consumer consent, the relevant legal construct is any consumer confusion at the time the firm obtains consumer consent, regardless whether the firm had a hand in producing that confusion.

- Clear and Conspicuous Disclosure Requirements

Under U.S. law, many properties of products and services, and financial products in particular, must be clearly and conspicuously disclosed to consumers. For example, under both the Truth in Lending Act and the Truth in Savings Act, a long list of items must be “clearly and conspicuously” disclosed to consumers.82 A variety of information in consumer auto leases must be set forth “in a clear and conspicuous manner.”83 Any deposit-taking institution that lacks federal deposit insurance must include “a clear and conspicuous notice” disclosing that fact in its advertising.84 In both financial and nonfinancial contracts, U.S. consumers can waive most of their legal rights, but courts usually require those waivers to be clear and/or conspicuous.85

No consistent test for what is “clear and conspicuous” has been developed. Sometimes, a regulator will issue a “model form” that, when used, gives the firm a safe harbor from liability for a violation of a clear and conspicuous disclosure requirement. The reliance on model forms treats clarity and conspicuousness as an

---

78 Negative option plans are contractual terms that require consumers to take an affirmative action to avoid being repeatedly provided with and charged for goods or services. These are also called continuity plans, when they automatically follow an initial membership period that the consumer intended to obtain. See 16 C.F.R. § 310.2(u) (defining “negative option feature” as “an offer or agreement to sell or provide any goods or services, a provision under which the customer’s silence or failure to take an affirmative action to reject goods or services or to cancel the agreement is interpreted by the seller as acceptance of the offer”).
79 16 C.F.R. § 310.4 (a)(7). The Telemarketing Sales Rule was promulgated pursuant to the Telemarketing and Consumer Fraud and Abuse Prevention Act, which requires the Federal Trade Commission and the Securities and Exchange Commission to “prescribe rules prohibiting deceptive telemarketing acts or practices and other abusive telemarketing acts or practices.” 15 U.S.C. § 6102(a) & (d).
82 12 C.F.R. § 1026.5; 12 C.F.R. § 1030.8.
84 12 C.F.R. § 1009.4.
inherent quality of particular text, devoid of context and irrespective of audience. Statutes and courts also sometimes treat clarity and conspicuousness requirements as if they were fixed characteristics, identifiable by font size, audio volume, or other simplistic tests. With this formalistic approach, evidence of actual consumer confusion is legally irrelevant.

Courts have moved toward more functional approaches for these requirements. Most courts agree that a conspicuous disclosure is one that a reasonable consumer would notice and a clear disclosure is one that a reasonable consumer would understand. A disclosure cannot be conspicuous if it is hidden in legal notices that consumers do not read, among many lines of small print text, or behind website links that a reasonable consumer would not click. A disclosure cannot be clear if it is ambiguous, meaning reasonable consumers can interpret it in more than one way.

Even under a functional approach, some courts treat clarity and conspicuousness as a question of law not a question of fact, thus calling for a judge’s assessment as to what reasonable consumers would notice and understand rather than for evidence of whether actual consumers notice and understand the disclosures. In 2002, one court opined: “We decide conspicuousness as a matter of law. This is not because judges are experts at graphic design but because subjecting conspicuousness to fact-finding would introduce too much uncertainty into the [firm’s] drafting process.”

Today, given the extent of firm testing and continual revision of their own consumer-facing communications, this reasoning is woefully outdated. Firms use eye-tracking, link-clicking, and mouse-hovering studies to assess where on their websites their customers are looking. Therefore, firms are increasingly likely

---

86 Many of these forms are developed through consumer testing but in artificial lab conditions or in field tests using cooperating firms that are aware of the testing. See Lauren E. Willis, The Consumer Financial Protection Bureau and the Quest for Consumer Comprehension, 3 Russell Sage Foundation Journal of the Social Sciences 74 (2017).

87 See, e.g., 13 Pa. Cons. Stat. § 1201 (“A printed heading in capitals (as: NONNEGOTIABLE BILL OF LADING) is conspicuous. Language in the body of a form is conspicuous if it is in larger or other contrasting type or color.”); Cal. Bus. & Prof. Code § 17601 (“Clear and conspicuous” or “clear and conspicuously” means in larger type than the surrounding text, or in contrasting type, font, or color to the surrounding text of the same size, or set off from the surrounding text of the same size by symbols or other marks, in a manner that clearly calls attention to the language. In the case of an audio disclosure, “clear and conspicuous” and “clearly and conspicuously” means in a volume and cadence sufficient to be readily audible and understandable.); Corbett v. Firstline Security, Inc., 687 F. Supp. 2d 124, 132 (E.D.N.Y. 2009) (holding that contract terms gave consumers sufficient notice because, among other things, the terms were written in capital letters, a legible font size, and “primarily” in language understandable to a layperson). But see In re Bassett, 285 F.3d 882, 886 (9th Cir. 2002) (“Lawyers who think their caps lock keys are instant ‘make conspicuous’ buttons are deluded. In determining whether a term is conspicuous, we look at more than formatting.”).

88 See, e.g., Channell v. Citicorp National Services, Inc., 89 F.3d 379 (7th Cir. 1996) (“clear and conspicuous manner” means visible, with manner referring to mode of presentation, not degree of comprehension).

89 See, e.g., Barrer v. Chase Bank USA, N.A., 566 F.3d 883, 892 (9th Cir. 2009) (holding that Truth in Lending Act disclosures were not clear and conspicuous); U.C.C. § 1–201(b)(10) (defining conspicuous as “so written, displayed, or presented that a reasonable person against which it is to operate ought to have noticed it.”); Lundquist v Security Pacific Automotive Financial Services Corp., 993 F.2d 11, 15 (2nd Cir. 1993) (holding that to be “clear,” car lease provisions must be understandable to an ordinary consumer). But see Applebaum v Nissan Motor Acceptance Corp., 226 F.3d 214, 220–22 (3rd Cir. 2000) (holding that a disclosure under the Consumer Leasing Act is “clear” so long as an expert could understand it; a consumer need not understand it).

90 Rubio v. Capital One Bank, 613 F.3d 1195 (9th Cir. 2010) (applying Truth in Lending Act).


92 In re Bassett, 285 F.3d 882, 885 (9th Cir. 2002).

to know whether the disclosures on their websites are sufficiently conspicuous to be noticed by their customer base. In one pending case, the firm had performed eye-tracking and click-testing studies for marketing purposes prior to litigation, and had determined that consumers paid attention to the headline price in the company’s advertisements, and not the fine-print disclosures. Firms use comprehension testing to assess consumer understanding of instructions and safety warnings. They use rapid iterative testing of consumers’ responses to different designs of websites and apps, pricing structures, sales and return processes, and even products and services themselves. Thus, many firms already have information about what their customers’ eyes are drawn to and what their customers know, and can experiment with disclosure methods to increase clarity and conspicuousness. Thus, any uncertainty in the drafting process is waning.

Regardless of whether courts treat the issue as one of law or fact, courts are increasingly willing to look beyond their own reactions to determine whether a disclosure is clear and conspicuous.

- U.S. courts look to empirical studies performed outside of the litigation. For example, a court faced with the claim that the use of the term “fixed” interest rate in a credit card solicitation was not “clear” considered empirical studies that had been commissioned from a market research firm by the Federal Reserve Board. Those studies found that consumers commonly harbored the belief that a “fixed” interest rate cannot be changed so long as the cardholder uses that credit card. Accordingly, the court held that the issuer’s use of the term “fixed” to mean something different (the issuer used “fixed” to mean “not fluctuating with an index”) in promotional materials was not clear as required by the Truth in Lending Act.

- Another court looked beyond the document distributed by the lender to comply with the Truth in Lending Act’s disclosure mandate and found other communications from the lender that could have rendered the mandated disclosure, in context, unclear or inconspicuous.

Looking at the entirety of the context in which consumers receive disclosures to determine whether these disclosures are clear and conspicuous is consonant with Federal Trade Commission commentary on the topic:

There is no set formula for a clear and conspicuous disclosure...

The ultimate test is not the size of the font or the location of the disclosure, although they are important considerations; the ultimate test is whether the information intended to be disclosed is actually conveyed to consumers.

---


97 Rubio v. Capital One Bank, 613 F.3d 1195, 1200 (9th Cir. 2010).

98 Roberts v. Fleet Bank (R.I.), 342 F.3d 260, 268 (3rd Cir. 2003), as amended (Oct. 21, 2003) (“[W]hile we recognize that the [Truth in Lending Act] only applies the “clear and conspicuous” standard to required disclosures, we conclude that the [Truth in Lending Act] permits us to consider materials outside of the [required disclosure document] in determining whether the credit issuer disclosed the required information clearly and conspicuously.

This ultimate test best meets the purposes of clear and conspicuous disclosure requirements. For example, the purpose of the disclosures required by the Truth in Lending Act is “to provide, especially to the inexperienced and uninformed consumer, a way to avoid the possibility of deception, misinformation, or at least an obliviousness to the true costs of a credit transaction.” Thus, the legally-relevant consumer confusion construct is confusion that is created by, or not cleared up by, the disclosure given by the firm.

Regardless of whether the definition adopted is formal or functional or whether treated as a question of law or fact, in all cases the law focuses on whether the firm’s communication of the terms that the statute requires to be disclosed is clear and conspicuous. Even if consumers are not confused because they already know the pertinent fact about the product or service, a firm has violated the statute if the firm fails to disclose that fact clearly and conspicuously.

<table>
<thead>
<tr>
<th>Type of Legal Claim</th>
<th>Relevant Consumer Confusion Construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>○ Unfair or abusive practices prohibitions</td>
<td>Confusion from which the firm profits (regardless of the source of that confusion); Mis-selling</td>
</tr>
<tr>
<td>○ Deceptive practices prohibitions</td>
<td></td>
</tr>
<tr>
<td>○ Informed consent requirements</td>
<td></td>
</tr>
<tr>
<td>○ Deceptive practices prohibitions</td>
<td>Confusion created by the firm (and from which the firm typically profits)</td>
</tr>
<tr>
<td>○ Clear and conspicuous disclosure requirements</td>
<td></td>
</tr>
<tr>
<td>○ Misleading representations prohibitions, including false competitor advertising claims</td>
<td></td>
</tr>
</tbody>
</table>

2. The Presence of Inaccurate Beliefs or the Absence of Knowledge

**KEY CONCEPT:** Consumer confusion that is relevant to U.S. consumer protection and fair competition law can include both affirmatively inaccurate beliefs and a lack of knowledge about key facts relating to a transaction.

Consumers can be confused due to affirmatively inaccurate beliefs about key facts or due to a lack of knowledge about key facts. The line between the two can be tricky, and courts considering claims involving consumer confusion typically do not make the distinction explicit. However, whether only the presence of inaccurate beliefs, or also the absence of knowledge, is the legally-relevant confusion construct depends on the law in question.

**Informed consent laws** aim directly at lack of knowledge but also require that the knowledge be accurate. Take as an example an investor who incorrectly believes that his or her investment advisor has no conflict of interest. This investor cannot give informed consent to the conflict, just as an investor who has no idea whether

---

101 See, e.g., Rubio v. Capital One Bank, 613 F.3d 1195, 1199 (9th Cir. 2010) (noting that the Truth in Lending Act’s disclosure requirements demand “absolute compliance by creditors”).
his or her investment advisor has a conflict of interest cannot give informed consent to that conflict. Both the presence of incorrect beliefs and the absence of knowledge are types of actionable confusion here, because in both conditions a consumer is insufficiently informed for the firm to obtain the consumer’s informed consent.

Laws requiring clear and conspicuous disclosures likewise aim directly at the absence of consumer knowledge rather than the presence of inaccurate consumer beliefs. The stated purpose of the Truth-in-Lending Act is “to assure a meaningful disclosure of credit terms so that the consumer will be able to compare more readily the various credit terms available to him and avoid the uninformed use of credit” in short, to avoid consumer ignorance of key facts. However, courts have further held that the information the disclosure is required to convey must be accurate, so as to avoid not only the uninformed but also the misinformed use of credit.

❖ For example, the Truth-in-Lending Act requires specified key facts about the cost of credit to be disclosed clearly and conspicuously. In several cases, courts have held that misleading disclosures violate the Act. In a case involving a credit card issuer, the court explained, “[b]ecause the purpose of the TILA is to assure meaningful disclosures, the [credit card] issuer must not only disclose the required terms, it must do so accurately.”

Unfair practices prohibitions, as explained above, are intended to prevent firms from “creating or taking advantage of an obstacle to the free exercise of consumer decision making.” Both the presence of a false belief about a key fact and the absence of knowledge about a key fact could be such an obstacle. As also noted above, among the prototypical examples of unfair firm behavior identified by the Federal Trade Commission are “withholding or failing to generate critical price or performance data ... leaving buyers with insufficient information for informed comparisons.” This suggests that the relevant confusion construct in unfairness cases includes a lack of knowledge.

Abusive practices prohibitions proscribe “taking unreasonable advantage of (A) a lack of understanding on the part of the consumer of the material risks, costs, or conditions of the product or service [or] (B) the inability of the consumer to protect the interests of the consumer in selecting or using [the product or service].” Both the presence of a false belief about a key fact and the absence of knowledge about a key fact are a “lack of understanding” and could prevent consumers from protecting their own interests when engaging

---


104 Roberts v. Fleet Bank (R.I.), 342 F.3d 260, 266 (3d Cir. 2003), as amended (Oct. 21, 2003). See also Jenkins v. Landmark Mortgage Corp. of Virginia, 696 F. Supp. 1089, 1094 (W.D. Va. 1988) (holding that when a Truth-in-Lending Act disclosure was accompanied by a cover letter inaccurately contradicting the disclosure, the disclosure did not meet the “clear and conspicuous” requirements of the Act).


106 Id.

107 12 U.S.C. § 5531(d)(2). Some state consumer protection statutes retain an old common law intent requirement, defining taking advantage of consumer ignorance as an unconscionable practice only if the business knew or should have known the consumer’s ignorance. See, e.g., Ohio Rev. Code § 1345.03 (listing as factor to consider is determining whether act is unconscionable, in violation of state UDAAP law, “[w]hether the supplier has knowingly taken advantage of the inability of the consumer reasonably to protect the consumer’s interests because of the consumer’s ... ignorance”). Today, given the amount of consumer testing that firms perform, it is arguable that firms should know when their customers are ignorant or confused about a key fact pertaining to the product or service.
in financial transactions. The consumer confusion that abusive practices lawsuits are intended to address thus can include both the presence of a false belief and the absence of knowledge.

In contrast, laws prohibiting false competitor advertising prohibit firms from creating affirmatively inaccurate consumer beliefs, and do not address the absence of consumer knowledge. For example, one credit card issuer probably would not be able to sue another issuer for unfair competition on the grounds that the second issuer’s marketing did not inform customers how much they would have to pay to transfer a balance to the second card. On the other hand, if the second issuer’s marketing affirmatively led consumers to believe that they would not be charged any fee for transferring a balance to that second issuer’s card, the first issuer could have a false competitor advertising claim against the second issuer.

Distinguishing the presence of inaccurate beliefs from the absence of knowledge can be challenging because when consumers do not know a key fact about a product or service, they often assume something about that fact, thereby forming an affirmative inaccurate belief. Most deception by omission cases are of this type:

❖ For example, a credit card issuer that disclosed all fees except a 1% currency conversion fee in monthly statements (the fee was disclosed only in the cardholder agreement), was found to have engaged in a deceptive practice. The court found that reasonable consumers would expect all fees to appear together, and that without knowledge of the fee, consumers reasonably assumed there was no such fee, and were thereby deceived.108

In other situations, experimental research suggests that consumers have a tendency to treat missing information as indicating that the product or service is of average or modal quality on the attribute for which information is missing.109 Take, as a hypothetical example, a consumer comparing five investment products, one of which clearly carries a particular fee, three of which clearly do not charge that particular type of fee, and one of which has information about the fee hidden in fine print such that the consumer does not know whether it carries the fee. The experimental evidence suggests that in this situation, some consumers might assume that the investment product about which he or she lacks knowledge does not charge this fee, because most of the products do not charge the fee.

---


109 See, e.g., Sunita Sah & Daniel Read, Research: Missing Product Information Doesn’t Bother Consumers as Much as It Should, Harvard Business Review (Sept. 28, 2017) (providing experimental subjects with reports about ratings patients have given doctors on a variety of dimensions, with some doctors’ scores on trustworthiness randomly missing, and finding that subjects on average treat doctors with missing information as being similar to the average doctor for whom there is complete information); Nikolos Gurney & George Loewensteine, Filling in the Blanks: What Customers Assume about Potentially Valuable but Missing Information, Oct. 2017, Figures 2, 3 & 4, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3050641 (finding evidence that subjects on average treat restaurants that randomly lack a sanitary inspection grade as being about average, and reminding subjects about the dangers of foodborne illness did not change the results).
### Mapping Claims to Constructs

<table>
<thead>
<tr>
<th>Type of Legal Claim</th>
<th>Relevant Consumer Confusion Construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>o Misleading representations prohibitions, including false competitor advertising claims</td>
<td>Presence of inaccurate beliefs</td>
</tr>
<tr>
<td>o Informed consent requirements</td>
<td>Presence of inaccurate beliefs and Absence of knowledge</td>
</tr>
<tr>
<td>o Clear and conspicuous disclosure requirements</td>
<td></td>
</tr>
<tr>
<td>o Unfair or abusive practices prohibitions</td>
<td></td>
</tr>
<tr>
<td>o Deceptive practices prohibitions</td>
<td></td>
</tr>
</tbody>
</table>

3. **The Persistence, Precision, and Depth of Consumer Knowledge**

**KEY CONCEPTS:** When assessing consumer confusion, the consumer knowledge that is legally relevant is knowledge of key transaction facts. To meet the purposes of consumer protection and fair competition law, this consumer knowledge ideally should be of sufficient persistence, precision, and depth to facilitate autonomous consumer decisions such that collectively, consumer choices can induce increased substantive market competition. The ideal persistence, precision, and depth is highly context-specific. However, decreasing consumer confusion is likely to be valuable even if consumer knowledge does not reach the ideal.

For the purposes of U.S. consumer protection and fair competition law, what does it mean to be confused or deceived about a transaction? For all claims, to be legally actionable the confusion must pertain to key facts regarding the transaction, including the cost to the consumer, limitations on benefits the transaction will provide, and whether the consumer has engaged in a transaction at all.

How long must a consumer’s knowledge of key facts persist in his or her memory, how precise must a consumer’s knowledge of these facts be, and how deeply must a consumer understand these facts for that consumer to be deemed not confused? The required persistence, precision, and depth *ought* to be that which is needed for the consumer to choose according to his or her own needs and desires and to engage in comparison shopping that can fuel substantive competition. This kind of understanding is not sufficient to ensure autonomous choices by consumers and substantive competition between firms, but it is necessary. However, the consumer confusion construct used today in U.S. litigation tends to demand a weaker, less precise, and more superficial level of knowledge.

The *persistence* with which knowledge about key facts must remain encoded and accessible in a consumer’s memory so as to be able to use that information to make marketplace decisions is likely to vary by context. The complexity of the service or product, the difficulty of mapping its features to the consumer’s personal needs, preferences, and constraints, and the manner in which consumers shop for it could all affect the necessary length of time. When consumers can compare products or services side-by-side using just a couple of key facts, information needs to be attended to and encoded to be used, but it need not be strongly encoded. For example, the price of using one online money transfer service to send money within the U.S. can

---

110 Roughly speaking, courts in the U.S. refer to key facts as “material facts.” Cf. Federal Trade Commission v. Cyberspace.Com, L.L.C., 453 F.3d 1196, 1201 (9th Cir. 2006) (holding that a misrepresentation is material if it makes it more likely that the consumer will choose the product being advertised).

be compared to the price of another online money transfer service in less than a minute. The price knowledge that firms need to convey to consumers for these types of services to facilitate autonomous consumer decisions and promote price competition probably does not need to last long.112

For a more complex product, for which many key facts must be compared, or when shopping takes place over a longer period of time, information might need to be more strongly encoded in memory. For example, shopping for a life insurance policy today often requires the consideration of many price structures, benefits, and limitations on benefits, and might take place over a period of time while the consumer compares information from disparate sources. The consumer might need to “sleep on it” to think over the uncertainties in his or her life and the lives of his or her beneficiaries so as to make the best decision for his or her circumstances and values.113 Knowledge of various key facts might therefore need to be encoded more strongly in memory to be used in comparison shopping for a life insurance policy.

Survey and usability experiments and copy and usability tests performed for litigation purposes in the U.S. usually do not require subjects to display knowledge that is very firmly encoded in memory. Typically, subjects are asked questions about a stimulus immediately after having read or viewed it, and sometimes the stimulus remains available to consult. In contrast, surveys of actual customers about their past transactions will often require the information to be more durably encoded in memory for the respondent to recall or recognize the information accurately.

The precision with which key facts must be known to best facilitate decisions that increase individual welfare and promote substantive competition will also vary with context, depending on, e.g., the individual’s situation and the structure of the market for the product or service. For example, a consumer with tighter budgetary constraints would likely need more precise knowledge of prices than a consumer who has slack in his or her budget. When the market for a product or service is not very competitive, even fairly imprecise consumer knowledge might increase competition. Where the market is already competitive with respect to the key fact in question, more precise knowledge is likely to be needed for consumer choices to lead to greater competition.

The knowledge treated as sufficient in research performed for litigation tends to be fairly imprecise. For example, when asked about a stimulus, subjects are typically not asked to recall a precise fee amount but rather whether the product or service does or does not carry a particular type of fee. Similarly, answers to open-ended questions about fee amounts are usually considered correct if the answer is close to the actual fee, and “correct” closed-end answer choices are often gross ranges of dollar amounts within which the fee could have fallen.114

Depth of knowledge refers to the degree to which a consumer understands key facts well enough to (a) understand how those key facts are likely to play out over time, given the consumer’s personal situation, and (b) compare the key facts about products and services offered in the marketplace. Knowledge of the consumer’s likely personal future experience with the product or service is likely to be more valuable for

112 Short-term memory, which lasts for between a few seconds and a minute, can retain four to five items or chunks of information. See, e.g., N. Cowan, The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity, 24 Behavioral and Brain Sciences 87 (2001), doi:10.1017/S0140525X01003922. However, there is evidence that consumer decision quality declines markedly if more than two or three attributes are considered. See Pete Lunn et al., PRICE Lab: An Investigation of Consumers’ Capabilities with Complex Products (May 2016), https://www.esri.ie/pubs/BKMNEX306.pdf.


encouraging substantive competition than abstract knowledge of key terms.\textsuperscript{115} For example, knowing that a credit insurance product covers car loan payments during temporary periods of unemployment due to disability but excludes disability due to pre-existing conditions is not very helpful if the consumer does not know that the back pain he or she reported to a doctor years ago will be considered a pre-existing condition excluding coverage for temporary unemployment caused by re-injuring his or her back in the future.\textsuperscript{116} Abstract knowledge of fee amounts for two investment products that have different fee structures is less useful than being able to compare the likely costs of those two products.

Respondents and subjects in social-scientific research performed for litigation are generally not asked to demonstrate a deep understanding of a stimulus. For example, subjects typically would not be asked to use a credit card issuer’s marketing materials to compute the total fees they would pay if they transferred their balance on an existing credit card to the issuer’s card. Instead, subjects would more likely be asked whether there is a fee at all for transferring a balance. Similarly, research performed to demonstrate legally-actionable consumer confusion does not ask subjects to use disclosures or marketing to choose among products or services, meaning that a consumer who knows, for example, all fees, but cannot use this information to select among offerings, is nevertheless deemed not confused.

Yet, it seems plausible that even superficial understanding, provided it is understanding of the truly relevant metric, would improve consumer autonomy and substantive competition. For example, during the U.S. mortgage bubble of the mid-2000s, many adjustable rate mortgage borrowers either did not know how much their payments could change or substantially underestimated how much they could rise. These same borrowers were disproportionately likely to have mortgages with potential monthly payment increases that could require the borrowers to pay more than an additional 8% of their monthly income toward their mortgages.\textsuperscript{117} Confusion audits to assess borrower knowledge of whether and by how much their payments could increase might have helped some consumers steer clear of unaffordable mortgages.

In sum, the persistence, precision, and depth of knowledge that should be required to demonstrate that any particular customer is not confused cannot be determined in the abstract. Instead, these dimensions of the consumer confusion construct that should be assessed for any particular legal claim will require judgment calls based on contextual factors.

\textbf{B. Caps on Customer or Consumer Confusion}

Not all confusion can be avoided, just as pollutant emissions and food contaminants cannot be entirely eliminated, at least not at an acceptable cost to society. For every product or service, there will always be some confused customers, some of whom believe the transaction has lower costs or higher benefits than it truly does, and others of whom believe it has higher costs or lower benefits than it truly does.

If sufficiently few customers are confused, and all customers are sold the same product or service on the same conditions and at the same price, savvy customers can, to the extent that the product is suitable for the savvy and confused customers alike, push the market toward fair competition over price and quality. Although these predicates for savvy consumers to push the market to low prices and high quality are disappearing—in the twenty-first century, individual consumers are with increasing frequency being sold different products and


services and at different prices—U.S. courts have generally not required firms to eliminate all consumer confusion or to avoid all transactions that profit from consumers’ confusion.

### 1. Explicit Qualitative Caps and Implicit Quantitative Caps

**KEY CONCEPT:** When enforcement authorities bring actions under consumer protection and fair competition laws, U.S. courts tend to find liability if a “significant minority” of consumers are confused, which courts interpret as meaning at least roughly 10% of consumers.

When U.S. courts consider the prevalence of confusion among likely consumers of a particular product or service or actual customers of a particular firm, the explicit standards for acceptable levels of confusion are all qualitative. For deceptive practices “a significant minority of reasonable consumers” must be shown likely to be confused for the firm practice to be actionable.\(^\text{118}\) When a firm is targeting a specific audience, confusion levels must be ascertained with respect to that group.\(^\text{119}\) For misleading advertising claims brought by one firm against another, the qualitative threshold is similar; implicit false competitor advertising claims are actionable when a “substantial segment” or “not insubstantial number” of the intended audience is likely to be misled.\(^\text{120}\)

Generally, courts have avoided explicit quantification of consumer confusion rate caps. Nonetheless, and without voicing a clear legal or normative defense for their numbers, the qualitative standards have been translated by courts into implicit quantitative caps on consumer confusion. In deceptive practices cases, the lowest levels of consumer confusion found actionable hover around 10% of the relevant audience.\(^\text{121}\) In firm-to-firm unfair advertising cases the implicit confusion rate cap is also about 10%.\(^\text{122}\) In other words, if evidence indicates that over 90% of the relevant consumers understand a key fact about a product or service, courts usually find that the plaintiff has not shown the defendant to have engaged in deceptive practices or false competitor advertising.

---


\(^{119}\) See, e.g., Lavie v. Procter & Gamble Co., 105 Cal. App. 4th 496, 508 (2003) (finding liability for deceptive practices under California law when “a significant portion of the consuming public or of targeted consumers acting reasonably under the circumstances, could be misled”); id. at 506–07 (“[U]nless the advertisement targets a particular disadvantaged or vulnerable group, it is judged by the effect it would have on a reasonable consumer.”).

\(^{120}\) See, e.g., Muzikowski v. Paramount Pictures Corp., 477 F.3d 899, 907 (7th Cir. 2007); McNeil Lab v. American Home Products, 501 F. Supp. 517, 528 (S.D.N.Y. 1980). For trademark matters, the same general formulation applies; a mark must be likely to confuse an “appreciable number” or “significant percentage” of the relevant consumer population. See, e.g., IDV North America, Inc. v. S & M Brands, Inc., 26 F. Supp. 2d 815, 829 (E.D. Va. 1998); Entrepreneur Media, Inc. v. Smith, 279 F.3d 1135, 1151 (9th Cir. 2002).

\(^{121}\) See, e.g., In the Matter of Telebrands Corp., 140 F.T.C. 278, 325 (2005) (10.5% confusion rate sufficient to show deceptive practices); ITT Continental Baking Co. 83 F.T.C. 865 (1973) (9% to 10% consumer confusion rate sufficient to demonstrate deceptive practices); Firestone Tire & Rubber Co. v. Federal Trade Commission, 481 F.2d 246, 249 (6th Cir. 1973) (upholding finding of deception where survey experiment indicated that 10% of the “buying public” was misled by the defendant’s marketing).

\(^{122}\) See, e.g., Mutual of Omaha Insurance Co. v. Novak, 836 F.2d 397, 400 (8th Cir. 1987) (survey experiment evidence indicating that 10% of subjects took away false messages from an advertisement (i.e., the “net takeaway” of a false message was 10%) was enough to support finding that false claim was communicated to consumers in business-to-business false competitor advertising case); Goya Foods, Inc. v. Condal Distributors, Inc., 732 F. Supp. 453, 456-57 (S.D.N.Y. 1990) (net takeaway of 9% justified finding false implied claim was made in false competitor advertising case). But see Millennium Importing Co. v. Sidney Frank Importing Co., No. CIV.03-5141 JRT/FLN, 2004 WL 1447915, at *11 (D. Minn. June 11, 2004) (11% confusion rate was not enough to demonstrate false competitor advertising). Consumer confusion rates found sufficient to demonstrate trademark infringement have been similar. Below 10%, courts have been divided as to whether the evidence demonstrates infringement. Compare Grotian v Steinway & Sons, 365 F. Supp. 707, 716 (S.D.N.Y. 1973) (finding 7.7% consumer confusion rate sufficient for trademark infringement) and Coca-Cola v. Tropicana Products, Inc., 690 F.2d 312, 317 (2nd Cir. 1982) (7.5% confusion rate sufficient), with Sara Lee Corp. v. Kayser-Roth Corp., 81 F.3d 455,467 n.15 (4th Cir. 1996) (empirical evidence of a consumer confusion rate under 10% warrants a finding that the alleged infringing mark does not create a likelihood of consumer confusion).
2. Barriers to Quantification

Explicit quantitative rules for caps on acceptable levels of consumer confusion are eschewed by U.S. courts for several reasons, including, e.g.:

- Some methodologies courts use to assess likelihood of confusion (e.g., unaided judicial judgment, firm behavior and intent evidence, expert facial analyses, and usability inspections), do not produce quantitative results. Courts demand quantitative evidence for firm-to-firm unfair advertising claims but do not always require it in other cases involving consumer confusion.

- Even in those cases in which quantitative evidence of consumer confusion is required, this evidence is weighed along with various qualitative factors in assessing legal liability.

- Different quantitative confusion measurement methods, no matter how perfectly performed, produce different results. A customer survey, for example, can capture both confusion caused by the firm and pre-existing confusion from which the firm may profit. A survey experiment, in contrast, attempts to control for pre-existing confusion.

- So too, question design will affect measured confusion rates. For example, closed-ended questions measuring consumer confusion might generate fewer “do not know” answers and high frequencies of guessing, because respondents sometimes guess even when they do not know the answer. On the other hand, open-ended questions, keeping the underlying confusion rate constant, can produce answers that are less complete because known information is not brought to mind.\(^{123}\)

- Measured confusion rates will also depend on what persistence, precision, and depth of knowledge is required of a respondent or subject to count his or her response as knowledgeable and not confused, discussed above. For example, if any answer other than the precise amount of a particular fee is coded as demonstrating the respondent to be confused, the measured rate of confusion will likely be higher than if less precision is demanded.

- It is impossible to measure confusion perfectly. All measurement methodologies are imperfect. Sometimes even the direction in which the imperfection will bias results is unclear. For example, when a study employs a nonprobability (convenience) sample—a routine practice of firm marketing departments—the results are “rough indicators” rather than “precise estimates.”\(^{124}\) In other situations, the direction in which a flaw is likely to bias results is ascertainable, but the magnitude of that bias is unclear.\(^{125}\)

---

\(^{123}\) See below for further discussion of open-ended and closed-ended survey questions.

\(^{124}\) See Shari Seidman Diamond, *Reference Guide on Survey Research*, in Reference Manual on Scientific Evidence 359, 383 (Federal Judicial Center: 3rd ed. 2011), (“[Q]uantitative values computed from [convenience/nonprobability] samples (e.g., percentage of respondents indicating confusion) should be viewed as rough indicators rather than as precise quantitative estimates.”).

\(^{125}\) For example, in one trademark case 67% of survey experiment subjects were confused by the non-holder’s use of the holder’s trademark, but due to multiple flaws in the research, the court determined that these results would not support a finding of infringement. Winner International, L.L.C. v. Omori Enterprises, Inc., 60 F. Supp. 2d 62, 70-71 (E.D.N.Y. 1999). In contrast, in another case the court held that a nearly identical confusion rate (70%), even after discounting the survey experiment’s results due to methodological flaws, remained sufficient evidence to support a finding of infringement. Academy of Motion Picture Arts and Sciences v. Creative House Promotions, Inc., 944 F.2d 1446, 1455-56 (9th Cir. 1991).
3. Setting Customer Confusion Rate Caps

**Key Concept:** The simplest way to set customer confusion rate caps is to start at a high level of confusion and reduce it over time, or to set a graduated staircase of caps and corresponding sanctions.

Even if empirical evidence could provide perfectly accurate and precise numbers, the normative question of how much consumer confusion ought to be tolerated would still be a challenge to answer. Some possible ways to set caps would include cost-benefit analysis, market-structure analysis, competitor performance levels, regulator experiments, or an incremental approach.

**Cost-benefit analysis** would be a normatively strong method for setting confusion caps, were it possible to perform. The regulator would determine the costs and benefits of eliminating the confusion of the marginal customer and set the cap where benefits just outweigh costs. However, these costs and benefits are themselves difficult to quantify. For costs, one would need to play out each possible response of the firm to the confusion cap—changes in marketing, sales processes, or product or service terms and structure—as the true cost is only the least expensive of these. Measurement challenges would plague estimates of benefits as well, including not only the benefits of increased competition but also increased consumer autonomy and trust in the marketplace.

**A market structure–based approach** would ask what proportion of customers (or segments of customers) must understand each key fact about the product or service so as to drive substantive competition. Is this a market in which understanding of costs and benefits by 10% of customers can drive firms to provide all customers the best possible product at the lowest possible price? The answer depends on, e.g.: whether savvy shoppers have the same needs, preferences, and use patterns as less knowledgeable consumers; whether the products and services are commodities or tailored; and whether firms can distinguish among consumers with varying confusion levels in a cost-effective manner. Economic modeling that took place before the modern marketing and product tailoring era estimated that if: (a) products are homogeneous; (b) sellers cannot distinguish between consumers on the basis of consumer knowledge levels; and (c) one-third of consumers in the market for the product in question understand product prices and engage in comparative shopping, then the market price will be driven to competitive levels (meaning the levels that would be produced if all consumers were fully informed).126

As firms engage in greater real-time hyper-targeted marketing, product and service design, and pricing, a larger proportion of consumers must comparison shop to make the marketplace competitive. If it is a market in which it is cost effective for firms to segment consumers and profit from uninformed market segments until 90% of customers know the key facts about the product or service, a 10% confusion cap (requiring 90% of each firm’s customers to understand the key facts about the product or service) makes sense. Although the entire marketplace is not there yet, the profitability of using big data to carefully price-discriminate seems likely to push all firms in this direction. In one analysis, price discrimination based on the crude demographic segmentation possible in the twentieth century pushed firm profits up by only about 0.3%; twenty-first-century

---

126 Alan Schwartz & Louis L. Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 University of Pennsylvania Law Review 630, 655 (1979). But see R. Ted Cruz & Jeffrey J. Hinck, *Not My Brother’s Keeper: The Inability of an Informed Minority to Correct for Imperfect Information*, 47 Hastings Law Journal 635, 672-74 (1996) (arguing, even before twenty-first century big data and micro-targeting of marketing, products, and pricing structures, that sellers were often able to differentiate among buyers and provide them with different terms, such that the “informed minority” of savvy shoppers did not set the terms that were offered to less knowledgeable buyers).
price personalization based on individual web-browsing data increased firm profits by 14.55% overall, with some individual consumers paying prices more than twice what others paid for the same product.127

Another potential system for setting customer confusion caps is through comparative analysis. How does the confusion rate of a firm’s customers stack up against other firms selling the same product? At the low end might be the median market-wide consumer confusion rate for the examined key product or service fact.128 A market median cap would tend to ratchet up over time, at least to the point at which the costs to firms of staying below the cap are less than the penalties incurred for exceeding the cap. However, entire markets could be performing so poorly, and ratcheting up might happen so slowly, that median-performance caps will only advance the goals of consumer law goals incrementally. More productively, a regulator might identify the customer confusion rate for top industry performers, those firms with relatively few customers who are confused about key product or service facts, and demand that all firms minimize their customers’ confusion to a similar extent.129 That other firms were able to reduce their customers’ confusion to this rate is an indication that it is possible to do so.

Here, care must be taken in defining each market. Some entire product types appear to be designed to be misleading. For example, the structure of payday loans appears intended to give the illusion that the cost of the loan to the borrower will be the dollar amount of the fee for one borrowing period (usually one pay period). However, many payday loan borrowers roll over their loans many times. At the time of loan origination, many payday borrowers may fail to appreciate the total costs they will pay after they repeatedly renew the loan. Therefore, customer knowledge of expected total borrowing cost could be low at most lenders, such that a confusion cap set by comparing different firms’ customers’ understanding of the payday loan credit product might not stimulate substantive price competition.130

In addition, if the cap is set by looking narrowly at the payday loan market, the system will not provide lenders an incentive to restructure their short-term loan product into a form in which their customers would better appreciate total borrowing costs, even though such a redesign might be the best way of producing consumer understanding and price competition. An alternative would be to define each market by the purpose the consumer seeks to meet in buying/using the product or service. For short-term borrowing needs, the U.S. market offers, e.g., credit cards, bank account overdraft, payday loans, and auto title loans. Arguably a comparatively-set confusion cap ought to be set with reference to these products collectively, so as to give firms an incentive to restructure their products to be the least confusing of these.131

---


128 This rate would need to be established through market-wide consumer testing. Cf. Amy Kapczynski & Ian Ayres, Innovation Sticks: The Limited Case for Penalizing Failures to Innovate, 81 University of Chicago Law Review 1781, 1830 (2015) (proposing a median market-wide performance measure be imposed on car manufacturers with respect to car-accident-related morbidity and mortality rates of occupants of their cars).

129 This is akin to environmental regulations, which are often set based on the performance achievable by the best technology currently available. See, e.g., Jody Freeman & Daniel A. Farber, Modular Environmental Regulation, 54 Duke Law Journal 795, 819 (2005).

130 Something similar may be true for debt relief services and credit insurance products, in that the products seem to be designed to be sold to consumers who do not understand the limitations on the benefits that these services and products provide. Confusion levels regarding benefits might be high at most sellers, such that using a comparative measure to set a customer confusion cap could be of limited utility.

131 Analogously, the EU has recognized the need to break from narrow product-specific disclosures and to instead require comparable disclosures for products that serve similar purposes for consumers. The EU has adopted a single uniform disclosure document—the “key information document” or “KID”—for packaged retail and insurance-based investment products, including, e.g., collective investment schemes, insurance-based investment products such as unit-linked life insurance products, and retail structured products. See Regulation (EU) No. 1286/2014 of the European Parliament and of the Council of 26 November 2014 on key information documents for packaged retail and insurance-based investment products (PRIIPs) [2014] OJ L352/1; EU Commission Staff Working Document on
An experimental approach to setting confusion caps would be to demand something close to what regulators can achieve in an experiment. This approach is similar to what happens in litigation today when the plaintiff performs a survey experiment, pitting test subjects’ understanding of a firm’s consumer-facing communications against a control group’s understanding of a control stimulus that attempts to hew as closely to the original as possible but subtracts the elements in the communication that are suspected of being misleading. Conceptually, the idea is that by subtracting the confusion rate produced in the control condition from the confusion rate produced when subjects are exposed to the firm’s materials, the resultant net confusion rate represents the confusion directly attributable to the misleading elements of the firm’s communications. As discussed above, courts tend to find a measured net confusion rate over 10% to be actionable. This 10% probably does not reflect a belief that each firm ought to be permitted to increase consumer confusion by 10 percentage points over the market-wide baseline. Rather, 10% might reflect a rough approximation of the highest estimate that is within the margin of error of zero.\textsuperscript{132}

A limitation of this experimental approach is that it assumes the product or service should be sold with its current design and features. Yet there are some financial product terms that are themselves so complex and counterintuitive that no method of disclosing those terms will lead to widespread consumer understanding. These would include double-cycle billing and applying credit card payments to cardholders’ lowest interest-rate balances first, as described in the Introduction to this Report. Confusion audits and caps are intended to drive at least some of these features out of the marketplace. An experimental approach retains these confusing features in the control condition, potentially depressing firm incentives to simplify terms or pricing even when doing so would be the best way to reduce customer confusion and increase substantive competition.

An incremental approach might be the simplest. The idea would be to set customer confusion caps initially at a high level, say 50%, while firms become accustomed to the system. Then, as firms become adept at compliance, periodically lower the cap down to the rate at which the estimate produced by customer surveys is statistically indistinguishable from zero, with some allowance for randomly-distributed customer confusion.

In conjunction with any of the above methods, a regulator could set a graduated schedule (staircase) of rewards or penalties to correspond to a range of customer confusion levels rather than a single cap. A staircase system could encourage firms to not only meet but exceed a particular level of customer understanding. It could also give industry laggards an incentive to lower their customer confusion rate, even if they might not be able to quickly meet a desired cap.

\textsuperscript{132} For a typical experiment with 200 subjects randomly drawn from the relevant population in each condition, at the 95% confidence level, the maximum margin of error is 10%. See, e.g., David A. Aaker et al., Marketing Research 402-431 (Wiley: 8th ed. 2004); Survey Money, Sample Size and Margin of Error Calculators, https://www.surveymonkey.com/mp/sample-size-calculator/ and https://www.surveymonkey.com/mp/margin-of-error-calculator/?utm_source1=mp&utm_source2=sample_size_calculator.
<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Strengths &amp; Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost-Benefit Analysis</strong></td>
<td>Set confusion cap where marginal benefits exceed marginal costs of the cap.</td>
<td>Theoretically strong but probably incalculable.</td>
</tr>
<tr>
<td><strong>Market-Structure Analysis</strong></td>
<td>Set confusion cap at level needed to incentivize substantive competition.</td>
<td>Theoretically strong but probably incalculable.</td>
</tr>
<tr>
<td><strong>Comparative</strong></td>
<td>Set confusion cap at the median rate achieved by competitors or the top quartile of performers in the product’s market.</td>
<td>Calculable but might not tap improvements that could be achieved through product or service design changes.</td>
</tr>
<tr>
<td><strong>Experiment-Based</strong></td>
<td>Set confusion cap at level achieved via communications that regulators design about the product or service.</td>
<td>Calculable but might not tap improvements that could be achieved through product or service design or pricing structure changes.</td>
</tr>
<tr>
<td><strong>Incremental</strong></td>
<td>Set high initial cap (perhaps 50%) and lower it over time.</td>
<td>Easy to administer but difficult to select schedule of caps to use over time.</td>
</tr>
<tr>
<td><strong>Staircase System</strong></td>
<td>A graduated system of penalties or rewards set in conjunction with one of the cap-setting methods above.</td>
<td>More likely to improve performance along the entire spectrum of firms but more difficult to administer.</td>
</tr>
</tbody>
</table>
III. Conventional Methodologies for Assessing Consumer Confusion in U.S. Legal Proceedings

Over time, a variety of epistemologies have been employed in U.S. legal proceedings to ascertain whether consumers are confused about the market transactions they have been offered or in which they have engaged. Until the latter part of the twentieth century, consumer perceptions were generally determined through the use of unaided judicial judgment, testimony from consumer witnesses, and evidence of firm or consumer behavior that sheds light on the question of consumer confusion. This Part of this Report discusses these conventional methodologies for assessing consumer confusion and catalogues their principal strengths and weaknesses. Although all of these methodologies continue to be employed in legal cases today, albeit sometimes *sub rosa*, they are often accompanied by social-scientific forms of evidence.

<table>
<thead>
<tr>
<th>Type of Evidence</th>
<th>Appropriate Construct</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaided Judicial Judgment</td>
<td>o Confusion created by the firm</td>
<td>o Economical</td>
<td>o Lack of population validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Unbiased by customer or firm Interests</td>
<td>o Lack of ecological validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>o Looming impracticability</td>
</tr>
<tr>
<td>Customer Witness Testimony</td>
<td>o Confusion created by the firm and/or from which the firm profits</td>
<td>o Strong ecological validity</td>
<td>o Weak population validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Can produce useful qualitative data</td>
<td>o Potential for measurement error</td>
</tr>
<tr>
<td>Firm Behavior &amp; Intent Evidence</td>
<td>o Confusion created by the firm and/or from which the firm profits</td>
<td>o Strong ecological validity</td>
<td>o Potentially ambiguous</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Economical</td>
<td>o Looming impracticability</td>
</tr>
<tr>
<td>Customer Behavior &amp; Dissatisfaction Evidence</td>
<td>o Confusion created by the firm and/or from which the firm profits</td>
<td>o Strong ecological validity</td>
<td>o Weak population validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Can produce useful qualitative data</td>
<td>o Potentially ambiguous</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>o Potential for measurement error</td>
</tr>
<tr>
<td>Customer Satisfaction Evidence</td>
<td>o None</td>
<td>o None</td>
<td>o No construct validity</td>
</tr>
</tbody>
</table>

1. Unaided Judicial Judgment\(^{133}\)

**Key Concept:** Unaided judicial judgment is economical but has serious validity problems and could become impracticable as the number of distinct communications and online interface designs produced by a single firm for selling a single product or service increases exponentially.

Traditionally, when a legal issue revolved around consumer perceptions and confusion, a court would take an advertisement, disclosure, product packaging, or other consumer-facing business communication and apply...
the court’s own judgment and common sense to ascertain the “net impression” the court believed these materials were likely to make upon consumers. In this approach, judges use their own intuitions and reactions to determine, for example, whether a document is deceptive, a disclosure is clear and conspicuous, or an online purchase process is misleading.

Courts have often framed this approach as an inquiry into the likely perceptions of “reasonable” or “average” consumers. Although the use of judicial judgment is thus couched as an “objective” way to identify violations of the law, unaided judicial judgment leans heavily on the subjective impressions of the particular judge.

Today, judges continue to employ their own judgment to analyze explicit claims made by firms and assess whether those claims are literally false or false by necessary implication. For claims that are implicit in consumer-facing materials produced by a firm, courts today often rely upon evidence from expert witnesses demonstrating what messages actual consumers are likely to receive.

- **Construct Validity**

For some types of legal claims, unaided judicial judgment presents a construct validity issue. Construct validity is the degree to which, even apart from methodological issues, research identifies or measures what it purports or intends to identify or measure. As explained above in this Report, different legal claims involve different consumer confusion constructs. In making an unaided judicial judgment, the judge examines only specific materials (e.g., advertisements, disclosures, sales scripts, and websites) and opines on whether those specific materials are likely to deceive consumers into inaccurate beliefs or effectively convey accurate information to consumers. This is the relevant construct for some deceptive practices claims, clear and conspicuous disclosure claims, and misleading representations/false competitor advertising claims.

When the litigation is about whether a firm has engaged in unfair or abusive practices or has obtained informed consumer consent, the misleading nature of particular materials can be too narrow an inquiry. Consumers receive information from multiple stimuli and from personal experiences, which collectively form consumer perceptions about goods and services. Profiting off of consumer confusion, regardless of the source of that confusion, is an unfair practice. When the legal question is whether a consumer has given informed consent, what matters is not whether a particular piece of text was clear and conspicuous but rather whether the consumer knew the pertinent information—from whatever source—prior to consenting to the transaction. Therefore, unaided judicial judgment is sometimes an inappropriate methodology for assessing consumer confusion related to these claims.

However, when it is reasonable to conclude that most consumers acted based on specific materials disseminated by the firm or that those materials were the only source from which consumers obtained key transaction facts, unaided judicial judgment regarding those firm’s materials can sometimes be an appropriate methodology. The firm’s materials alone cannot reveal whether consumers actually read the materials or understood them, so the unaided judicial analysis of those materials cannot demonstrate that the firm did not

---

134 In re National Credit Management Group, L.L.C., 21 F. Supp. 2d 424, 441 (D.N.J. 1998) (“In determining if deceptive claims ... are present, a court is not limited to express claims, but may also look to the overall net impression conveyed by the advertising and promotional statements of a defendant.”).

135 See, e.g., Strubel v. Capital One Bank (USA), N.A., 179 F. Supp. 3d 320, 325–26 (S.D.N.Y. 2016) (“The Court evaluates the adequacy of [Truth in Lending Act] disclosures from the vantage point of a hypothetical average consumer—a consumer who is neither particularly sophisticated nor particularly dense.”); Federal Trade Commission Policy Statement on Deception, 103 F.T.C. 174, 175 (1984) (“Whether a representation or an omission, the deceptive practice must be ‘likely to mislead a reasonable consumer acting reasonably under the circumstances.’”). As explained above, one statute, the Fair Debt Collection Practices Act, requires business communications to be reviewed from the standpoint of an “unsophisticated consumer” or the “least sophisticated consumer” rather than the “average” or “reasonable” consumer.
engage in unfair or abusive practices or did obtain informed consumer consent. Yet, if the firm’s materials failed to convey the information that a consumer must know to give informed consent, and the court concludes that it is more likely than not that consumers did not obtain this information from any other source, then unaided judicial judgment regarding the firm’s materials tests the correct construct for informed consent claims.

- **Strengths**

  The greatest virtue of unaided judicial judgment is how quickly and inexpensively it can be applied to resolve legal questions. A half-century ago, most firms each used only a handful of form documents, advertisements, and label designs. A judge could review these in short order and opine on whether the explicit or implicit claims or disclosures made therein were false, unclear, inconspicuous, or confusing. Further, unlike evidence generated by customers or firms, unaided judicial judgment is unlikely to be biased by customer or firm interests.

- **Weaknesses**

  The trouble with reliance on unaided judicial judgment to ascertain consumer perceptions is inaccuracy. Three major deficiencies, problems that in the language of the social sciences we might call a lack of “population validity,” a lack of “ecological validity,” and, for some types of legal claims, a lack of “construct validity,” undermine the approach. Further, the profusion of consumer-facing communications produced by firms may soon make unaided judicial judgment impracticable.

  **Lack of Population Validity.** Unaided judicial judgment lacks population validity because judges’ background knowledge, information-processing skills, and perceptual capacities often differ from the background knowledge, information-processing skills, and perceptual capacities of the consumers who engage in the transactions at issue. As U.S. Judge Jerome Frank famously explained 70 years ago in dissenting from his colleagues’ finding, based on their personal perceptions, that consumers were likely to be confused about whether a “Miss Seventeen” girdle was connected with “Seventeen” magazine:

  As neither the trial judge nor any member of this court is (or resembles) a teen-age girl or the mother or sister of such a girl, our judicial notice apparatus will not work well unless we feed it with information directly obtained from “teen-agers” or from their female relatives accustomed to shop for them.136

  What might confuse actual consumers of a particular product or service might not fool a judge, and what might fool a judge might not fool actual consumers of the product or service. For example, judges are likely to pay more attention to text and to have higher literacy and numeracy levels than average consumers.137 On the other hand, for many products, judges will not have personal experience or social knowledge that informs customers who do buy those products. A judge might think a girdle and a magazine were related products when a teenage girl would not. Moreover, modern marketing often features little text, and instead an

---

136 Triangle Publications, Inc. v. Rohrlich, 167 F.2d 969, 976 (2nd Cir. 1948) (Frank, J., dissenting). See also Williams v. Gerber Products Co., 552 F.3d 934 (9th Cir. 2008) (reversing trial court dismissal of deceptive advertising case because “the district court based its decision to grant the motion to dismiss solely on its own review of an example of the packaging”); Carter v. First National Collection Bureau, Inc., 135 F. Supp. 3d 565, 569 (S.D. Tex. 2015) (“As other courts have recognized ... district judges are not good proxies for the unsophisticated consumer whose interests the [Fair Debt Collection Practices Act] protects....”).

137 Vanishingly few consumers read all the text accompanying transactions. Yannis Bakos et al., *Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts*, 43 Journal of Legal Studies 1 (2014), https://doi.org/10.1086/674424 (finding that fewer than two out of every thousand consumers choose to access software license agreements, and those one or two who look at the contracts read no more than a small portion of the text, on average).
abundance of imagery or scenarios that might convey a message to customers but not to judges outside the target market.\(^{138}\)

Courts have attempted to address the population validity problem through use of a standard reflecting the perceptions of a “vulnerable” consumer or an “average” consumer to which the particular product or service was targeted.\(^{139}\) However, without access to the thought processes of vulnerable or other particular subgroups of consumers, judges can only conjecture how these consumers perceive the pertinent advertising, disclosures, websites, etc. A judge is likely to have a personal reaction to the evidence in mind and, due to anchoring and adjustment bias, is unlikely to make sufficient adjustments to that reaction in speculating on the vulnerable or average consumer’s reaction.\(^{140}\)

Moreover, consumers’ perceptions are often unexpected, even for professional marketers who have greater expertise in this area than judges. Firms today, and financial services firms in particular, rarely rely on their own judgment when making decisions that depend on understanding consumer perceptions and instead gather empirical evidence about actual consumer perceptions. As the then-Chief Executive Officer of Capital One, a large credit card issuer, has explained, the credit card business is “a scientific laboratory where every decision about product design, marketing, channels of communication, credit lines, customer selection, collection policies and cross-selling [can] be subjected to systematic testing using thousands of experiments.”\(^{141}\)

**Lack of Ecological Validity.** Unaided judicial judgment lacks ecological validity because the judge examines the advertising, packaging, or other consumer-facing material in a decontextualized setting. Salesperson body language and tone of voice, the flow of a website purchase process, and the distractions of everyday life are difficult if not impossible to recreate in evidence submitted to the court. What might deceive, confuse, or miss the notice of a consumer focused on completing a transaction in the hurly burly marketplace might not deceive, confuse, or miss the notice of a judge focused on assessing the matter in the quietude of chambers.

**Looming Impracticability.** Online marketing is moving from a mass advertising model in which only a few formats were used, to a micro-targeted model in which the marketing copy itself varies based on what the firm knows about the consumer generally (e.g., education level, age) and at that point in time (e.g., whether the consumer is inebriated or stressed) and the environment surrounding the consumer (e.g., time of day, recent high-profile events).\(^{142}\) Further, online marketing materials can take a host of formats, from whole websites to display advertisements to “native advertising” (e.g., videos, social media posts, or “news” stories designed to

---


139 See, e.g., Lavie v. Procter & Gamble Co., 105 Cal. App. 4th 496, 512 (2003) (“Where the advertising or practice is targeted to a particular group or type of consumers, either more sophisticated or less sophisticated than the ordinary consumer, the question of whether it is misleading to the public will be viewed from the vantage point of members of the targeted group, not others to whom it is not primarily directed.”).


appear as if they are created by consumers or journalists rather than by the firm). The designs of all of these can evolve very rapidly, using machine learning to discover which design features produce the consumer responses desired by the firm. Moreover, not only is the design of the marketing moving to micro-targeted, real-time forms but attributes of products and services will soon follow. Machine-driven dynamic price optimization, by which different consumers are offered different prices or pricing structures, is already happening in the financial services marketplace.

This presents a looming impracticibility problem for unaided judicial judgment (and for survey and usability experiments, copy and usability tests, and expert facial analyses and usability inspections, discussed below). The following series of cases illustrates the problem:

➢ In a 2010 case, the Federal Trade Commission, in its adjudicative capacity, analyzed in detail each of 36 different advertisements distributed by a single firm for a single product and determined whether consumers were likely to take away false messages from each advertisement. In the course of its opinion, the Commission explained that evidence regarding consumers’ interpretations of one set of the firm’s advertisements (advertisements on billboards) would not demonstrate whether the firm’s other (print) advertisements conveyed deceptive claims. The Commission explained that although the print versions did contain the same headlines and imagery that appeared on the billboards, the print advertisements contained additional text that might change the message consumers received. The Commission thus implied that every design difference in firm communications requires a separate analysis.

➢ Five years later, in a case brought by the Federal Trade Commission against a firm alleging a failure to adequately disclose certain facts about the transaction in the firm’s online marketing, the firm explained that it “engaged in ‘multivariate testing,’ in which live websites were repeatedly changed to determine which changes increased purchases,” resulting in “hundreds (or possibly thousands) of different [web]pages.” The firm argued that the analysis of 125 of these websites by the Commission’s expert was not sufficient to demonstrate anything about the defendants’ other websites. The court never directly addressed this issue but proceeded to painstakingly review the dozens of websites the Commission had submitted to the court. Eventually, after several rounds of hearings and a bench trial,

---

143 See, e.g., Sami Main, New Study Shows that the Number of Native Ad Buyers Increased by 74% in Just One Year, AdWeek (July 10, 2017), http://www.adweek.com/digital/new-study-shows-that-the-number-of-native-ad-buyers-increased-by-74-in-just-one-year/; Andrew Graham, Native Advertising, Explained: Cannes Edition, AdWeek (June 23, 2015), http://www.adweek.com/digital/native-advertising-explained-cannes-edition/ ("Native advertising is the practice of making an ad on a platform resemble whatever content the user is there to see or read.").


the court found each one to be deceptive and/or to have failed to make certain required disclosures in a clear and conspicuous manner.\textsuperscript{147}

➢ In a case pending as of this writing (June 2018), the defendant has asserted that a survey experiment conducted on one print advertisement and a usability experiment conducted on one website configuration cannot demonstrate that it engaged in deceptive practices or failed to provide consumers with clear and conspicuous disclosures because during the relevant time period, the defendant disseminated “more than 40,000 unique print ads, hundreds of banner ads, dozens of TV commercials, and many different iterations of [a] continually evolving website.”\textsuperscript{148} The court has not yet ruled on this issue.

Accordingly, for a judge to apply his or her unaided judgment to a firm’s consumer-facing communications will become increasingly burdensome as the number of unique websites, videos, advertisements, disclosures, and other communications multiplies. Courts might be willing to analyze a representative sample of these communications rather than requiring the plaintiff to present each one to the court.\textsuperscript{149} However, it is not clear how a representative sample could be compiled. Courts also might be willing to perform an in-depth analysis of a handful of websites, and then scan the others to determine whether they convey “the same information to [consumers] in the same language,” rendering the websites “essentially identical.”\textsuperscript{150} But even analyzing a dozen multi-page websites is burdensome, and human scanning of 40,000 print advertisements would take a very long time, detracting from the primary strength of unaided judicial judgment—how quickly and inexpensively it can be performed.

2. Customer Witness Testimony

\textbf{KEY CONCEPT:} Customer witness testimony has weak population validity for identifying the presence or absence of confusion about key product or service facts but can provide useful qualitative data.

One time-honored practice is for courts to allow customers to testify about their own experiences and perceptions. The party alleging a violation of consumer protection or fair competition law trots out a parade of consumer witnesses, each of whom testifies about being misled and confused. Firms respond with testimony from their own satisfied, knowledgeable customers.

For example, in a case regarding informed client consent to an investment adviser’s conflict of interest, testimony of the clients was introduced to demonstrate whether the clients understood the conflict of interest, such that they could have given informed consent to the transactions.\textsuperscript{151} Consumer testimony about their own experiences and perceptions is admissible not to demonstrate what the firm actually did or said but to


demonstrate what the consumers knew or believed or what impression the firm’s words and actions left on consumers.\textsuperscript{152}

- **Strengths**

  **Strong Ecological Validity.** What actual customers understood about a transaction is the relevant construct for determining whether a firm engaged in unfair or abusive practices or obtained informed consent from its customers. The problem of a judge guessing at what consumers would have thought, without knowing the full history and context that shaped the consumers’ beliefs about their transactions, is eliminated. Ecological validity is strong for this kind of evidence.

  **Can Produce Useful Qualitative Data.** Importantly, customer witnesses can provide detailed qualitative data that can shed light on the sources and parameters of customers’ confusion in a particular case. Follow-up and cross-examination questions can be asked to assure that the court understands what the witness means by his or her testimony. For example, in the investment adviser case just mentioned, the issue was whether the adviser’s clients gave their informed consent to the profits the adviser made by buying securities and then reselling them to clients from her own account. Some customer witnesses testified that they could calculate these profits from the documents the adviser gave them, implying that their consent was informed. When given the documents on the witness stand and asked to explain how they would calculate those profits, the clients realized they could not do so.\textsuperscript{153}

- **Weaknesses**

  **Weak Population Validity.** Unless a very high proportion of a firm’s customers testify, attorney-selected samples of customers will typically lack the representativeness that would be needed to generalize about the perceptions of a firm’s customer population as a whole. Therefore, population validity is problematic. Moreover, the testifying customers are likely to not only be attorney-selected but also self-selected, and those who volunteer their experiences are likely to differ, possibly in relevant ways, from those who do not.

  **Potential for Measurement Error.** Further, customer witnesses may well have a personal stake in the outcome of the case. For unfair and deceptive practices cases, customer witnesses who testify to being duped may be seeking compensation. By the same token, customer witnesses who testify to having understood the key facts of the transactions they engaged in with the firm may be seeking to vindicate their own decisions to transact with the firm. Social scientists might call this a problem of measurement error, in that the data collected—the testimony—is potentially biased.

  Another potential source of measurement error is the delay between the transaction and the testimony. Memories may have faded, such that even customers who were not confused at the time of the transaction could be confused at the time of testimony. The opposite could also be true; customers who were confused at the time of the transaction could have become informed in the meanwhile, perhaps even through the litigation. People have poor insight into how they come to know things,\textsuperscript{154} and so customer witnesses could have difficulty parsing what they knew at the time of the transaction from what they know at the time of trial.

**The Use of Customer Witness Testimony Today**

\textsuperscript{152} See, e.g., Federal Trade Commission v. John Beck Amazing Profits, L.L.C., 865 F. Supp. 2d 1052, 1081 (C.D. Cal. 2012) (explaining that consumers’ testimony does not establish the content of the business’s marketing materials; rather, the testimony is “relevant to show the consumer’s understanding of the statements made in the [marketing] materials”).


Despite weak population validity and the potential for measurement error, customer witness testimony can help the court understand the context surrounding transactions and the prior experiences of consumers that shape their perceptions of key facts about the products and services at issue in the case. The testimony can also perform a function akin to pretesting surveys with a small sample of consumers or debriefing experimental subjects. Witnesses can be asked questions that were posed in, for example, a survey of a representative sample of customers that is already in evidence, and then asked follow-up questions to demonstrate how the survey answers might have been interpreted or misinterpreted by the survey respondents. For these and other reasons, courts today routinely take testimony from consumer witnesses.

3. Firm Behavior and Intent Evidence

**KEY CONCEPT:** Firm intent with respect to customer confusion is typically not crucial to claims brought under U.S. consumer protection and fair competition laws. Nonetheless, firm behavior and intent evidence can be revealing. However, this type of evidence could become increasingly difficult to obtain as computer algorithms direct more firm marketing and sales activity with little human involvement.

Generally speaking, consumer protection and fair competition laws create strict liability claims, meaning that violations do not depend on the firm’s intent. However, evidence of intent can be introduced and can relieve the plaintiff of the need to produce empirical evidence of consumer confusion. If a firm intended to deceive consumers or to convey a false claim to consumers, the courts will generally assume it was successful in doing so.

Sometimes a firm will take actions that reveal its knowledge or intent that consumers’ confusion will lead consumers to take actions favorable to the firm. This type of evidence might be found in a firm’s own records or in testimony from a firm’s current or former employees.

- For example, in one case alleging unfair and deceptive practices and Truth in Lending Act violations against a payday lender, it came to light that an employee had suggested providing consumers with a clearer explanation of loan terms. An email entered into evidence was used to show that management rejected the suggestion on the grounds that a clearer explanation would discourage consumers from taking out the loans.

---

155 The main exception would be securities laws, which prohibit false statements, misleading omissions, and fraudulent and manipulative practices only when the defendant acts recklessly or intentionally. 15 U.S.C. § 78j; 17 C.F.R. 240.10b-5.

156 See, e.g., CollegeNET, Inc. v. XAP Corp., 483 F. Supp. 2d 1058, 1065-66 (D. Or. 2007) (where business intended to deceive students about the privacy of their information so as to sell that information to lenders, competitor did not need to demonstrate that students were actually confused: “In the absence of actual evidence of deception, the court may presume consumers were deceived if the deception was intentional.... The defendant may rebut the presumption by presenting evidence that it did not succeed in the deception.”) (false competitor advertising); In re Telebrands Corp., 140 F.T.C. 278, 304 (2005) (showing that the business intended to make a particular claim is “powerful evidence that the alleged claim in fact was conveyed to consumers”) (unfair and deceptive practices violation); Johnson & Johnson*Merck Consumer Pharm. Co. v. SmithKline Beecham Corp., 960 F.2d 294 (2nd Cir. 1992) (intent to infringe is sufficient evidence of trademark infringement). But see Sims v. GC Services, L.P., 445 F.3d 959 (7th Cir. 2006) (“If debt collectors go to great lengths to produce confusing letters and attempt to deceive the recipients, their intent would not matter if the letters on their face contained the required notifications and would not confuse the unsophisticated consumer. Conversely, debt collectors might make every effort to make the letters clear and not confusing, yet if the letters would confuse the unsophisticated consumer and violate the statute, debt collectors would be held liable.”).

157 See, e.g., Davis v. Southern Bell Telephone & Telegraph Co., No. 89-2839-CIV-NESBITT, 1994 WL 912242, at *7 (S.D. Fla. Feb. 1, 1994) (testimony from the business’ former employees demonstrated that their telephone solicitations used unfair and deceptive practices to sell the service to consumers); Federal Trade Commission v. John Beck Amazing Profits, L.L.C., 865 F. Supp. 2d 1052, 1062 (C.D. Cal. 2012) (same). Although former employees could have a motive to put the business in a bad light, where this type of motivation is found the testimony is unlikely to be relied upon, absent corroborating evidentiary support.

Market research studies performed or commissioned by a firm in the ordinary course of its affairs can sometimes directly reveal a firm’s knowledge of consumer confusion.

- For example, in one case, the court relied on a firm-commissioned copy test in which respondents reacted to materials the firm used to sell existing customers a service. The copy test revealed that a diagram the firm was using left respondents confused about key facts consumers would need to know to evaluate the value of the service sold, and that elderly respondents could not read the materials due to small print. In its subsequent mailings, the firm continued to use the same diagram and font size and continued to omit the information consumers needed to correct their misimpressions about the value of the service. This evidence was used to demonstrate both that the firm’s materials were unfair and deceptive, and that the firm intentionally chose to use such materials, knowing that they were unfair and deceptive.159

Further, website usability testing is standard operating practice for many firms with websites today.

- In a pending case, the Federal Trade Commission has submitted evidence that the defendant, a telecom provider, performed an eye-tracking study, a click-test study, and multiple other website usability studies. Each study confirmed that consumers did not look at legal disclosures on the websites, did not notice links to disclosures that would have revealed key costs of the service, and did not scroll down to parts of webpages that do not appear immediately on the screen. In one of the usability studies, consumers “[f]elt it was difficult to track when price increases were happening and why” and felt that the purchase flow included examples of “bait and switch tactics.”160 The telecom continued to place key price and term information in legal disclosures, either at the bottom of websites and/or accessible only through clicking on links or info-hover icons (which the firm’s own studies found were rarely clicked or hovered over).

Even without a study, internal firm marketing communications can be revealing.

- In a case against Amazon regarding its distribution of “free” children’s internet game applications that produce revenue through sales made to players during the game (“in-app purchases”), the court relied in part on internal firm documents demonstrating that Amazon knew many customers (parents) did not understand in-app purchases, Amazon knew that children were initiating in-app purchases, and Amazon had itself noted “the disconnect between the account owner (e.g., parent) and the app user (e.g., child).” Based on these findings as well as an expert usability evaluation of the website by which Amazon distributed the games, the court determined that Amazon had committed an unfair practice by making sales to children playing video games using credit card information that their parents had previously provided to Amazon and without their parents’ knowledge or consent to the purchase.161

Another kind of firm behavior evidence that can demonstrate a firm’s intent to confuse consumers is the internet search keywords that the firm buys to determine which internet users receive the firm’s advertising.

- In a case alleging that a firm deceptively advertised its product to have a valuable benefit the product did not have, the court admitted evidence that the firm purchased internet advertising keywords


relevant to that benefit. This evidence contributed to the court’s ultimate finding that the firm deceived consumers about the benefits of the product.\textsuperscript{162}

Firm statements to investors have been used as evidence of a firm’s intent to deceive consumers.

\begin{itemize}
  \item In a case alleging that a car rental company deceived consumers into purchasing insurance (called a “loss damage waiver”) that the consumers did not need or want, the company’s 10-K Report (an annual report required by the Securities and Exchange Commission for all publicly-traded companies) stated:

  Legislation affecting the sale of loss damage waivers has been adopted in 25 states. These laws typically require notice to customers that the loss damage waiver may duplicate their own coverage or may not be necessary. Adoption of national or additional state legislation affecting or limiting the sale ... of loss damage waivers could result in the loss of ... revenue.

  As the expert witness in that case explained: “By [the firm’s] recognition that their profitability may be adversely affected by laws requiring them to give consumers notice that ‘the loss damage waiver may duplicate their own coverage or may not be necessary,’ they are recognizing that their firm model relies in part on an uninformed consumer that can be duped.”\textsuperscript{163}

\end{itemize}

To be clear, to violate U.S. law, a firm need not intend or even be aware that its practices are unfair, deceptive, or abusive, that consumers have not given it required informed consent, or that its disclosures are not clear and conspicuous. Firm behavior evidence is not admissible to demonstrate that the firm lacked intent or was unaware of the effects of its advertising, labeling, website design, or other consumer-facing materials, unless the party pursuing the firm has made intent or knowledge an issue.

\section*{Strengths}

\textit{Strong Ecological Validity.} The biggest strength of firm behavior and intent evidence is that it was not produced for the purposes of litigation, and therefore not crafted to favor the firm in the litigation.\textsuperscript{164} So long as the evidence was created for a business purpose unrelated to trying to make the business’s customers appear less confused than they are, it has strong ecological validity.

\textit{Economical.} In addition, to the extent that the firm-created evidence relieves the party bringing the case of the burden of obtaining surveys and other expert evidence, using that evidence to demonstrate consumer confusion can be economical.

\section*{Weaknesses}

\textit{Looming Impracticability.} A few types of business behavior and knowledge evidence are common enough that they should be sought out. For example, studies performed for marketing purposes, website usability testing, and information about purchases of internet search keywords can be fertile sources.

\begin{itemize}
  \item \textsuperscript{162} Mullins v. Premier Nutrition Corp., 178 F. Supp. 3d 867, 879 (N.D. Cal. 2016) (in a case claiming that the business deceived consumers into thinking that its product would reduce arthritic joint pain (when the product had no demonstrated efficacy for arthritis), explaining: “When a Google user types in keywords, Google.com and its affiliates display advertisements for those companies that bid successfully for the [keywords].... [The business in this case] bid on the following [keywords]: ... “arthritis treatments,” “how to stop arthritis,” “supplement for arthritis,” ... and “arthritis cures.”).
  \item \textsuperscript{163} McKinnon v. Dollar Thrifty Automotive Group, Inc., Case 12:cv-04457-SC (N.D. Cal. 2015) (Declaration of Donald R. Lichtenstein, Ph.D., Dec. 30, 2014, at ¶ 38).
  \item \textsuperscript{164} This evidence is admissible despite the usual evidentiary prohibition in U.S. courts on hearsay, either under the business records exception, Federal Rule of Evidence 803(6), or as a party admission, Federal Rule of Evidence 801(2). See, e.g., Schering Corp. v. Pfizer, Inc., 189 F.3d 218, 238 (2nd Cir. 1999), as amended on rehearing (Sept. 29, 1999).
\end{itemize}
However, other types of evidence of a firm’s knowledge or intent can be difficult to locate, and are likely to become even more difficult to obtain in the future. Financial services firms in particular have turned to machine learning, which can continually tweak consumer-facing materials based on incoming data about click rates, sales, retentions, and the like. No intent may be discernable other than the intent to maximize various outcomes. To the extent that machine learning controls the development of online and offline firm communications, it is possible that no human would be directly involved in designing the content of these communications or determining the context in which these communications appear.

4. Customer Behavior and Dissatisfaction Evidence

**KEY CONCEPT:** Customer behavior evidence can have strong ecological validity for identifying customer confusion but can have weak population validity.

Sometimes consumers act in ways that reveal their confusion about key transaction facts. For example, if consumers open accounts or purchase memberships, evidence that these customers do not use the accounts or memberships suggests that they may not have realized they had opened the accounts or bought the memberships.\(^{165}\)

- For example, one case alleged that consumers were deceived into thinking they were buying an ordinary pre-paid general purpose credit card that could be used anywhere when the card actually could only be used to buy the defendant’s products. The court opined: “The fact that consumers did not purchase the defendant’s products after obtaining their credit cards, which they could use to buy only defendant’s products, suggests that they were actually deceived.”\(^{166}\)

Similarly, evidence that virtually all customers fail to complete programs for which they have paid is some evidence that the customers did not understand the requirements and limitations of the program before purchasing it.\(^{167}\)

Consumers also sometimes report fraudulent charges on their credit card statements and contact their credit card issuers, which then chargeback the firm for the amount.

- Several courts have found a chargeback rate that is higher than normal to be evidence that a firm’s customers were confused about the transaction when they engaged in it or did not know they were being charged by the firm until they saw their credit card statements.\(^{168}\)

- Similarly, that borrowers block one method by which a lender can electronically debit payments from their bank accounts has been taken as some evidence that these borrowers did not know that in the

---

\(^{165}\) See, e.g., Federal Trade Commission v. Cyberspace.Com, L.L.C., 453 F.3d 1196, 1201 (9th Cir. 2006) (“[The business] billed each of these consumers for a service that less than one percent of them ever attempted to use. It is reasonable to infer that most of the remaining 99 percent did not realize they had contracted for [the] service...”).

\(^{166}\) Federal Trade Commission v. USA Financial, L.L.C., 415 F. App’x 970, 973 (11th Cir. 2011).

\(^{167}\) See, e.g., Federal Trade Commission v. Capital Choice Consumer Credit, Inc., 2004 WL 5149998, at *35 (S.D. Fla. Feb. 20, 2004), affirmed, 157 F. App’x 248 (11th Cir. 2005) (evidence that less than one half of one percent of purchasers ever completed a program for which they had paid nearly $200 demonstrated that consumers were deceived as to the requirements of the program because “although it would be normal to expect that some consumers would fall behind in making their monthly payments and thus not complete the program, surely more than one half of one percent would have completed the program had they understood its requirements and limitations before paying the $189 or $199.95 fee”).

fine print of their loan documents they had agreed to allow the lender to use another method to electronically debit their bank accounts too.\textsuperscript{169}

Customers sometimes complain to firms directly or to third-parties (government agencies such as state attorney generals’ offices, the Federal Trade Commission, the Consumer Financial Protection Bureau, or nonprofit Better Business Bureaus) that the product or service they received was not what they believed they purchased or that the terms asserted by the firm post-sale differed from the terms customers were led to believe pre-sale. It is well-known that most dissatisfied customers do not complain, but evidence that more than a small proportion of a firm’ customers have done so has been accepted by courts as some evidence of unfair and deceptive practices.\textsuperscript{170}

- **Strengths**

  **Strong ecological Validity.** As with firm behavior and knowledge evidence, consumer behavior evidence is not created in anticipation of litigation, and therefore is not influenced by any interest the consumer may have in the litigation. It therefore has strong ecological validity.

  **Can Produce Useful Qualitative Data.** Sometimes consumer complaints are accompanied by narratives, akin to customer witness testimony. These narratives can provide useful qualitative data.

- **Weaknesses**

  **Weak Population Validity.** For customer complaint evidence, population validity is weak. Customers who complain differ from customers who do not, and complaint rates are likely to understate the incidence of customer confusion and unfair and deceptive firm practices.

  **Potentially Ambiguous.** Sometimes it is difficult to determine whether customer behavior evidence is evidence of customer confusion. In a recent case, an expert witness for a payday lender compared payoff behaviors of experienced and inexperienced borrowers and opined that his data showed their behaviors to be nearly identical. He further opined that this correlation demonstrated that borrowers were not deceived at origination, reasoning that experienced borrowers could not have been deceived due to their experience and that inexperienced borrowers would have had different payoff patterns than experienced borrowers if inexperienced borrowers had been deceived.\textsuperscript{171} The opposing expert opined, among other things, that any correlation in behavior between experienced and inexperienced borrowers did not support the conclusion that borrowers were not deceived, because, e.g., research about consumer psychology and consumer learning demonstrated that experienced borrowers would not necessarily have learned from experience and therefore could also have been deceived when they took out their loans.\textsuperscript{172}


\textsuperscript{170} See, e.g., State v. Vertrue, Inc., 2010 WL 1652448, affirmed 834 N.W.2d 12 (Iowa 2013) (“[T]he patterns that emerge from the larger pool of self-generated consumer complaints received by the [Better Business Bureau] are highly probative of whether Defendants’ practices have a tendency or capacity to deceive, or are unfair.”); “The Court need not, and does not, credit each and every factual allegation put forward by each complaining consumer as necessarily true. Instead, the Court relies on the consumer complaints to the extent that a pattern emerges that corroborates other non-hearsay evidence the Court has received.”).


\textsuperscript{172} Federal Trade Commission v. AMG Services, Case 2:12-cv-00536-GMN-VCF, Document 940-7 (Rebuttal Expert Report of George Loewenstein, Ph.D.). This expert also explained that the first expert’s data analysis suffered from a well-known statistical error, and, properly analyzed, the data did not show any meaningful correlation between the payoff behaviors of the two groups.
5. A Note on Customer Satisfaction Evidence

**Key Concept:** Customer satisfaction evidence cannot demonstrate an absence of customer confusion about key product or service facts.

Defendants sometimes offer evidence of “satisfied” customers in an attempt to show that their products and services met customer expectations, and therefore their customers must not have been confused or deceived when they entered the transaction. This evidence takes two forms: customer satisfaction surveys and data about customer post-sale behavior, such as repeat purchases or a lack of complaints or returns. Although courts treat an unusually high volume of complaints, chargebacks, cancellations, or returns as some evidence of unfair and deceptive practices, an absence of complaints cannot be taken as evidence that firms did not violate the law. As courts have long held, “the existence of satisfied consumers does not constitute a defense” to unfair or deceptive practices claims.

Customer satisfaction evidence lacks construct validity with respect to claims involving consumer confusion. As courts have recognized:

- A low refund rate may not represent satisfaction. Even dissatisfied consumers may fail to exercise their right to a refund, because they think it not worth the trouble, because they feel guilty for having been deceived, because they credit the product’s ineffectiveness to their own failure to follow instructions, or for any one of a number of other reasons.

- In one unfair and deceptive practices case against a seller of “credit repair services” and secured credit cards, the court found it irrelevant that few consumers complained about the firm.

- In another case, a payday lender argued that borrowers who took out multiple loans must have been satisfied and therefore could not have been deceived. The court rejected this reasoning: “Respondents ... presented no evidence that consumers took additional loans because they were satisfied; repeat customers may simply have been caught in a cycle of debt and saw no other option than to take out multiple loans.”

Academic research in the consumer finance area supports judicial rejection of customer satisfaction data. For example, one study sent testers, who were blind to the purpose of the experiment, into the field to obtain advice from financial advisers. The experimenters gave the testers various fictitious financial profiles, and the testers reported back on the advice they received. Although the experimenters determined objectively that much of the advice the testers received was biased to favor the interests of the financial advisers, the testers overwhelmingly believed they had received excellent financial advice.

---


175 Federal Trade Commission v. Pantron I Corp., 33 F.3d 1088, 1093 (9th Cir. 1994).

176 See, e.g., In re National Credit Management Group, 21 F. Supp. 2d 424, 442, n. 31 (D.N.J. 1998) (rejecting argument where 280 of 200,000 clients complained about defendants’ misrepresentations and omissions regarding credit repair services and secured credit cards).


IV. SOCIAL-SCIENTIFIC METHODOLOGIES FOR ASSESSING CONSUMER CONFUSION IN U.S. LEGAL PROCEEDINGS

Mindful of the weaknesses in relying upon unaided judicial judgment, select customer witness testimony, and the happenstance of locating revealing business or consumer behavior evidence, U.S. courts have turned to evidence prepared by expert witnesses. This can take a variety of forms, including expert facial analyses or usability inspections of marketing materials or websites. For some types of legal claims, however, courts expect the parties to submit surveys, experiments, or tests performed by experts on a reasonably representative sample of actual customers or likely consumers.

For example, unless a communication conveys an explicit claim that is literally false or false by necessary implication, a business generally cannot successfully sue another business for false advertising without evidence about consumer beliefs gathered through a survey experiment:

To prove that an advertisement is impliedly false, the plaintiff must come forward with specific scientific survey evidence of consumer reaction. The Court’s own perception of whether or not the advertising is misleading is irrelevant and insufficient.

Another court explained:

A court may, of course, construe and parse the language of the [claim]. It may have personal reactions as to the defensibility or indefensibility of the deliberately manipulated words.... [B]ut the court’s reaction is at best not determinative and at worst irrelevant. The question in such cases is—what does the person to whom the advertisement is addressed find to be the message?

However, where the firm’s own marketing research demonstrates that it is misleading consumers, or evidence directly shows the firm’s intent to deceive, the opposing party is not required to produce its own survey.

The principal aim here has been to achieve population validity, in that expert judgment and tests of actual consumers can more accurately assess the perceptions of consumers than can a judge aided by only his or her imagination. Ecological validity remains imperfect for many forms of consumer confusion evidence, given that a consumer answering a questionnaire is not in the same environment and frame of mind as a consumer buying a product or service. But no piece of evidence is perfect. Results of properly-performed consumer surveys and experiments are sufficiently reliable to use in legal fora.

179 The Australian Full Federal Court has recognized the same. Arnotts Ltd. v. Trade Practices Commission (1990) 24 FeR 313, 360-1 (Full Federal Court) at 64 (noting, in the context of declaring survey evidence admissible to demonstrate consumer confusion in a trademark case, that “information is preferable to intuition”).

180 Several years ago, a number of studies of published judicial opinions in the U.S. suggested a declining reliance on survey evidence. However, a more thorough analysis of how attorneys use survey evidence has since demonstrated that surveys remain key to litigation-related assessments of consumer perception. Well before the stage at which the case might produce a published judicial opinion, cases with weak survey evidence are usually abandoned and cases with strong survey evidence are usually settled. See Shari Seidman Diamond & David J. Franklyn, Trademark Surveys: An Undulating Path, 29 Texas Law Review 2029 (2014).


This Part of this Report discusses the use of expert witnesses to provide social-scientific evidence to courts generally, and then describes each category of expert evidence and explains the advantages and disadvantages of using each to demonstrate consumer perceptions and confusion.

A. The Use of Expert-Generated Evidence

There are three issues that should be considered when any form of expert evidence is involved:

1. Admissibility
2. Qualifications of the expert
3. Objectivity of the expert

The following examines each of these in the context of the types of evidence used to assess consumer confusion.

1. Admissibility

**KEY CONCEPT:** Expert evidence regarding whether consumers are confused about the key facts regarding their transactions with firms is admissible in U.S. legal proceedings when the research is relevant, is the product of reliable principles and methods, and is performed with the rigor expected of experts in the field.

The U.S. Federal Rules of Evidence generally allow the admission of all relevant evidence, meaning evidence that “has any tendency to make a fact more or less probable than it would be without the evidence” where “the fact is of consequence in determining the action.” Relevant evidence can be excluded if the evidence “poses a danger of unfair prejudice, confusion of the issues, or misleading the jury,” but only if the probative value of the evidence is “substantially outweighed” by those dangers.

Regarding expert testimony, the Federal Rules provide that:

- The expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- The testimony is based on sufficient facts or data;
- The testimony is the product of reliable principles and methods; and
- The expert has reliably applied the principles and methods to the facts of the case.

The legal requirement that expert evidence be “the product of reliable principles and methods ... reliably applied” is a requirement that the evidence have what social scientists would call scientific “validity” (not scientific “reliability”).

Courts look to a host of factors in determining whether an expert’s analysis has sufficient “evidentiary reliability” (that is, scientific validity) to permit its admission into evidence. A nonexhaustive list of factors relevant to evidentiary reliability include: whether the methodology has been subjected to peer review and publication; the existence and maintenance of standards controlling the methodology; and whether the

---

184 Federal Rule of Evidence 401.
186 Federal Rule of Evidence 702.
methodology has achieved general acceptance in the relevant scientific or expert community. In assessing consumer confusion evidence the courts generally follow the best practices recognized by academics and business professionals working in the relevant field, whether it be psychology, marketing, human-computer interaction (HCI), statistics, or another field. When best practices in the field change, the courts follow these changes, albeit sometimes with a bit of delay. For example, as marketing professionals moved from mall intercept convenience samples to landline telephone random-digit-dialed samples to internet panel samples, the survey modalities used in expert witness research accepted by the courts evolved as well.

To be clear, parties proffering expert evidence do not need to “prove their case twice.” They “need not demonstrate to the judge by a preponderance of evidence that the assessments of their experts are correct,” only that their experts’ opinions are “reliable,” meaning that “the particular opinion is based on valid reasoning and [a valid] methodology.”

Many forms of expert evidence involve hearsay, which is inadmissible in U.S. legal proceedings as a general rule. U.S. courts have held that surveys that ask for respondents’ beliefs or reactions fall within the present sense impression exception to the hearsay rule. For example, in one case a survey of doctors asked respondents to relate the main messages they had received about a drug from a drug manufacturer’s representatives close in time after having received a marketing pitch about the drug from those representatives. As held by then-Judge (now U.S. Supreme Court Justice) Sonia Sotomayor, although the survey results would not be admissible to show what the marketing representatives said, the results were admissible to demonstrate the messages doctors took from the sales pitches, whether those messages were expressly stated or merely implied.

Ultimately, the job of courts is to decide where the line is between social science evidence that is sufficiently valid to aid the court and social science evidence that is not. A key touchstone here is whether the evidence is of a type and quality that “experts in the particular field would reasonably rely on ... in forming an opinion on the subject.” An expert must employ in litigation “the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.” Thus, social science evidence that uses methods similar to the methods used in studies that firms rely upon in making major financial decisions is usually admissible, assuming the evidence is relevant to the legal issue at hand.

This does not mean that a particular firm’s confidence in its own research is a reliable indicator of the quality of that specific research; rather, it is a recognition by the courts that if a firm is willing to wager substantial sums on research that contains flaws of some sort, the courts can rely on independent, unbiased research that contains the same flaws. Further, a firm that has relied upon flawed research for business

---

188 Id. at 593-94.
189 See, e.g., Clicks Billiards, Inc. v. Sixshooters, Inc., 251 F.3d 1252, 1262 (9th Cir. 2001) (to be admissible, survey evidence must be “conducted according to accepted principles”); Pittsburgh Press Club v. United States, 579 F.2d 751, 758 (3rd Cir. 1978) (survey is admissible “if the poll is conducted in accordance with generally accepted survey principles, and if the results are used in a statistically correct way”).
191 Federal Rule of Evidence 703.
purposes can hardly complain when an opposing party uses that same research to demonstrate that the firm’s practices confused consumers or profited from consumers’ confusion.\textsuperscript{195}

2. Qualifications

\textbf{KEY CONCEPTS:} To produce admissible social-scientific evidence, an expert must have sufficient knowledge and skills to perform the research properly and must have experience performing the same type of research or analysis in academic, business, or government settings. Experts qualified to perform the types of research admissible to demonstrate the presence or absence of consumer confusion can hail from a variety of formal fields of training.

Under the Federal Rules of Evidence, a witness may be qualified as an expert on the basis of “knowledge, skill, experience, training, or education.”\textsuperscript{196} A court should examine the “proposed expert’s full range of experience and training, not just his [or her] professional qualifications,”\textsuperscript{197} but the expert must have some training in or experience with the relevant type of evidence.\textsuperscript{198} Although there is a first time for everything, that first time should not take place when an expert is performing an analysis for litigation purposes.\textsuperscript{199}

As with an expert’s methodology, one strong indicator of an expert’s qualifications is whether that expert is relied upon by firms, government agencies, or other experts/academics. When an opposing party introduces consumer perception or confusion evidence created by a firm not for the purposes of litigation or a regulatory examination but in the ordinary course of its operations, the court is likely to assume without further inquiry that the personnel who performed the research were qualified to do so.\textsuperscript{200} If the expert had sufficient expertise for the firm to rely on, the expert presumptively has sufficient expertise for the court to rely upon.

There is no single field in which an expert must be qualified to provide evidence about actual or likely consumer confusion; different forms of evidence will require different types of expertise, and many of these types of expertise cut across various academic disciplines. The following includes examples of the kinds of

\textsuperscript{195} For example, in an administrative proceeding brought against a firm for deception, the judge relied in part upon a consumer copy test that had previously been performed by the firm to decide which advertisements to produce in final form and run. The firm had tested the advertisements “in a rough, unfinished form.” The judge noted that “[i]n the experience of [the company’s] marketing research department, the results obtained from copy testing rough versions of [the] ads provided an accurate measure of how those ads would communicate to consumers in finished form.” The judge accepted the results of the firm’s own tests as probative evidence that consumers were confused by the firm’s advertisements, reasoning that the firm itself “relied heavily upon these copy tests in making consumer research-based business decisions.” In the Matter of Novartis Corp., 127 F.T.C. 580, *617-18 (1999).


\textsuperscript{197} See, e.g., Valador v. HTC Corp., 2017 WL 1037589 (excluding expert who had no prior experience with survey experiments testing likelihood of consumer confusion).

\textsuperscript{198} See, e.g., LVL XIII Brands, Inc. v. Louis Vuitton Malletier S.A., 209 F. Supp. 3d 612 (S.D.N.Y. 2016) (excluding proffered expert on grounds that he “[d[id] not appear to have any training or experience measuring secondary meaning—whether through a belief marketing survey or otherwise”); Moore v. Weinstein Co., L.L.C., No. 3:09-CV-00166, 2012 WL 1884758 (M.D. Tenn. May 23, 2012), affirmed, 545 F. App’x 405 (6th Cir. 2013) (excluding proffered expert because prior to the litigation he had “never performed an analysis of consumer perception/confusion”). On the other hand, the expert need not have previously encountered precisely the same situation in conducting prior analyses. See, e.g., In re ConAgra Foods, Inc., 90 F. Supp. 3d 919, 954 (C.D. Cal. 2015), affirmed sub nom. Briseno v. ConAgra Foods, Inc., 844 F.3d 1121 (9th Cir. 2017) (“[T]he fact that conjoint analysis has not been used to isolate the exact attribute for which Howlett uses it here does not automatically render her methodology and conclusions unreliable. It is Howlett’s experience with conjoint analysis and the details of her proposed methodology that determine reliability.”).

\textsuperscript{200} See, e.g., In the Matter of Novartis Corp., 127 F.T.C. 580, 619-628 (1999).
qualifications that have been found sufficient for an expert to prepare or critique each of the main types of expert-generated consumer confusion evidence.

Courts have accepted holders of doctorates in fields ranging from statistics to economics to psychology, and even a marketing consultant with a master’s degree in journalism, as experts qualified to perform and critique customer and consumer surveys.\textsuperscript{201} The common thread among these experts is that all had performed surveys of consumers for businesses, government agencies, or academic purposes prior to testifying, and thus had experience performing this kind of research.

Among those who have performed survey and usability experiments and were accepted by the courts as experts include advanced degree holders in psychology, mathematical statistics, and business.\textsuperscript{202} What they all have in common is experience performing survey or usability experiments and similar consumer research for business, government, and/or academic purposes.

Experts who perform consumer copy tests exhibit a similar diversity of formal qualifications. For example, in one case, one expert who conducted a consumer copy test was a marketing professor with a doctorate in business, another had a master’s degree in business and was president of a large consumer research company, and a third expert was a marketing professor with a doctorate in social psychology.\textsuperscript{203} Again, the unifying theme here is experience performing similar research.

Firms rely upon experts from a variety of backgrounds to perform usability testing, although all might be considered experts in human-computer interaction (HCI). The HCI field itself is interdisciplinary, covering a range of topics relating to aspects of electronic displays with which consumers interact, from human visual perception (how the eye processes sensory data) to human cognition (how the brain interprets and classifies information). HCI experts can be found in a variety of academic fields, including computer science, human factors engineering, and psychology. One recent job listing for a usability tester/analyst listed as job requirements:

\begin{itemize}
  \item BA/BS or MA/MS in behavioral science or closely related discipline. Qualifying fields include but are not limited to: Human-Computer Interaction, Human Factors, Psychology (Cognitive or Experimental), Interface Design, or Computer Science.\textsuperscript{204}
\end{itemize}


\textsuperscript{203} In the Matter Of Novartis Corp., 127 F.T.C. 580 (1999).

While few such experts have yet testified in litigation, they are likely to do so with increasing frequency as more market transactions move online.\footnote{U.S. Department of Health & Human Services, \textit{Usability Testing}, Usability.Gov, https://www.usability.gov/how-to-and-tools/methods/usability-testing.html.}

Experts qualified to perform expert facial analyses or usability inspections have a variety of formal qualifications. But all have experience performing similar analyses for firms, government agencies, nonprofits, or in academic work. One co-author of this Report, a marketing professor with a doctorate in business who over many years had performed consumer research for government agencies and his own academic work using a variety of methodologies, recently performed an analysis of the clarity and conspicuousness of a payday lender’s online and print “Loan Agreement” with its customers.\footnote{Expert Report of Manoj Hastak, In the Matter of Integrity Advance, 2015-CPBP-0029, Document 087A (filed 05/10/2016), https://files.consumerfinance.gov/f/documents/087-A_Expert_Report.pdf.} Another co-author of this Report, who at the time held a masters in information management and systems and had extensive experience performing usability inspections for industry, has provided expert opinions regarding ways in which the designs of firms’ websites and software applications were likely to mislead consumers.\footnote{Expert Report of Jennifer King, Federal Trade Commission v. Commerce Planet, Inc., 2011 WL 8611471 (C.D. Cal.) (filed March 7, 2011). The other report was submitted to the court by the Federal Trade Commission in In Re: Federal Trade Commission v. Amazon, Inc., 71 F. Supp. 3d 1158 (W.D. Wash. 2014), but has been filed under seal.} At least one expert with a doctorate in psychology and another expert whose highest degree is a juris doctorate (law degree) have also testified regarding electronic interface usability and deceptive design.\footnote{Declaration of David J. Franklyn, Federal Trade Commission v. DirectTV Case 3:15-cv-01129-HSG Document 189-19 (filed 10/06/16) (J.D. highest degree); Dorothy Atkins, \textit{DirectTV Website Ignored HHS Usability Rules, Expert Says}, Law360, Aug. 16, 2017, https://www.law360.com/articles/954826/directv-website-ignored-hhs-usability-rules-expert-says (Dr. Theo Mandel, Ph.D. in psychology).}

3. Ensuring Objectivity

**KEY CONCEPT:** Transparency sufficient for the opposing party and the court to fully assess expert evidence is the best way to protect against expert bias or error.

Bias on the part of an expert witness can undermine an analysis conducted by that expert. Where the “expert” is closely related to the firm, such as an officer or employee of the firm, the court is likely to exclude or discount the evidence entirely.211 These individuals are almost invariably not truly experts in any case, in that they lack the necessary credentials to perform the analysis and/or they have little or no experience with the type of analysis performed. Offering such an individual as an “expert” is rare.

A more serious concern is the “hired gun” problem – when qualified experts allow their expert opinions to be compromised by their affiliation with their clients.212 One rather pessimistic judge has declared:

> The battle of experts ... is frequently unedifying.... Many experts are willing for a generous (and sometimes for a modest) fee to bend their science in the direction from which their fee is coming. The constraints that the market in consultant services for lawyers places on this sort of behavior are weak, as shown by the fact that both experts in this case were hired and, we have no doubt, generously remunerated even though both have been criticized in previous judicial opinions. The judicial constraints on tendentious expert testimony are inherently weak because judges ... lack training or experience in the relevant fields of expert knowledge. But that is the system we have.213

However, when the methodology applied is one in which judgment calls must be made—a category into which much expert testimony regarding consumer perceptions falls—it is inevitable that courts will be faced with dueling expert opinions. Judicial frustration is about analyses that appear to produce contradictory results, and while conscious and subconscious expert witness bias certainly exists, the source of difference in the outcomes of expert facial analyses can usually be identified from an objective standpoint. Thus, courts rarely find bias alone to be a reason to exclude expert evidence from consideration and will instead examine the expert’s analysis and conclusions and weight the evidence according to its quality.214

Moreover, the adversary system provides its own check on the problem; the opposing side can present evidence of an expert’s bias and can critique the methodological soundness of the expert’s report. To this end, transparency is crucial. An expert performing survey research should provide sufficient information to enable a complete review of the methodology he or she employed and the data the research produced. The Council of American Survey Research Organizations has created a useful checklist in this regard.215

---

211 Rulings on Certain Motions In Limine at pp. 2-3, M2 Software, Inc. v. Madacy Entertainment, Case No. CV-00-2853-AHM (RZx) (C.D. Cal.), affirmed, 421 F.3d 1073, 1087 (9th Cir. 2005) ("expert" report was excluded because the expert was M2’s "CEO, president and only employee," who had never before designed a consumer survey).


213 Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Ltd. Partnership, 34 F.3d 410, 415 (7th Cir. 1994).

214 See, e.g., Bowers v. Norfolk Southern Corp., 537 F. Supp. 2d 1343, 1354-59 (M.D. Ga. 2007), affirmed, 300 F. App’x 700 (11th Cir. 2008) (finding that a physician proffered as an expert was a “hired gun” who spent a significant amount of his time serving as an expert witness rather than seeing patients, yet nonetheless carefully examining the physician’s opinions and rejecting them upon determining that the so-called expert had not used the same degree of care in preparing his expert report as he did in his practice as a physician and his opinions lacked sufficient scientific basis).

Finally, when skeptical of the expert opinions on both sides, courts are not helpless. A federal court in the U.S. can appoint its own expert to perform an independent analysis (e.g., a survey or consumer copy test).\textsuperscript{216} To assure that this expert is genuinely neutral, the court can ask the parties’ experts to jointly select an independent expert.\textsuperscript{217}

B. The Four Main Types of Social-Scientific Evidence for Assessing Consumer Confusion

The four main social-scientific types of evidence used in U.S. legal proceedings today to demonstrate consumer confusion are:

1. consumer and customer surveys;
2. survey experiments and usability experiments (with a control condition);
3. consumer copy tests and usability tests (without a control condition); and
4. expert facial analyses and usability inspections.

These labels have evolved over time and are used differently in different academic fields. Consumer or customer surveys, survey experiments, and copy tests are all frequently referred to by U.S. courts as “consumer surveys.” As used in this Report, the distinction is that consumer or customer surveys aim to discover existing knowledge and beliefs held by consumers or by the customers of a particular firm, whereas survey experiments and copy tests are designed to elicit consumer reactions to a stimulus, such as an advertisement or website. Consumer and customer surveys are evidence of actual consumer or customer confusion, no matter the source of that confusion. Survey experiments and copy tests are evidence as to whether the particular materials tested confuse consumers.

Sometimes the term “copy test” is used to include survey experiments. Both attempt to measure consumer perceptions of a stimulus, usually a firm’s marketing materials. Survey experiments use a control condition, randomizing subjects into a treatment group and a control group. As the term is used in this Report, copy tests show a stimulus to participants and ask for their reactions without comparing their responses to others who are shown a control stimulus. Sometimes courts call survey experiments “controlled copy tests.”

Usability experiments, usability tests, and usability inspections are ways to examine how the design of an electronic interface, such as a website or software application, communicates and interacts with users (consumers). In the field, terminology is inconsistent, and “usability evaluation” can refer to all or just one of these.\textsuperscript{218}

In this Report, a usability experiment refers to randomizing subjects into a treatment group and a control group and comparing the perceptions they draw from, and their ability to perform tasks on, a firm’s electronic interface and a control interface. This is analogous to a survey experiment.\textsuperscript{219} Usability testing refers to providing consumer subjects with an electronic interface and examining the reactions they take away from the

\textsuperscript{216} Federal Rule of Evidence 706. See also DeKoven v. Plaza Associates, 599 F.3d 578, 583 (7th Cir. 2010).


\textsuperscript{218} See, e.g., David Lucius, Marketing Madness: Website Usability Testing, Digital River, https://www.digitalriver.com/marketing-madness-website-usability-testing/ (employing “usability testing” to refer to an experiment in which subject interactions with and beliefs about two websites are compared); Jeff Sauro, 5 Types of Usability Tests, August 11, 2015, https://measuringu.com/five-types-usability/ (referring to what this Report calls usability experiments as "competitive usability testing").

\textsuperscript{219} See, e.g., Expert Report of Tülin Erdem, Ph.D., Federal Trade Commission v. DirecTV, Case 4:15-cv-01129-HSG, Document 261-3 (Filed 01/03/17) (reporting results of a survey experiment performed on one of the defendant’s print advertisements and a usability experiment performed on one of the defendant’s websites).
interface and their ability to use the interface to perform desired or assigned tasks.\textsuperscript{220} This is analogous to consumer copy tests. A usability inspection is performed by experts who examine the user interface and, applying their expertise in user reactions and behavior, assess whether consumers would be likely to be confused or have difficulty using the interface to perform desired tasks.\textsuperscript{221} This is analogous to an expert facial analysis of offline business-to-consumer communications.\textsuperscript{222}

This Report uses the four categories as further described below to permit consistent comparisons and meaningful analyses of the courts’ treatment of these four types of evidence. The following discusses the legal, methodological, and practical strengths and weaknesses of these various types of research in the context of assessing consumer or customer actual or likely confusion. For further information on methodological issues, the reader is advised to consult the \textit{Reference Guide on Survey Research}, which is the primary reference used by federal judges in the U.S. for assessing the scientific quality of surveys and survey experiments.\textsuperscript{223}


\textsuperscript{222} See, e.g., Federal Trade Commission v. Commerce Planet, Inc., 878 F. Supp. 2d 1048, 1071 (C.D. Cal. 2012), affirmed in part, vacated in part, remanded, 815 F.3d 593 (9th Cir. 2016), and affirmed in part, 642 F. App’x 680 (9th Cir. 2016) (“The Court finds that a usability inspection, with its emphasis on user perception and comprehension of the information presented to them on a webpage, is consonant with a “net impression” test ... which turns on a facial examination of the relevant marketing materials.”).

### Social-Scientific Evidence for Assessing Consumer Confusion

<table>
<thead>
<tr>
<th>Type of Evidence</th>
<th>Appropriate Construct</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer &amp; Customer Surveys (without a stimulus)</td>
<td>o Confusion from which the firm profits&lt;br&gt;o Confusion of a firm’s actual customers, created by the firm</td>
<td>o Potential for strong population validity</td>
<td>o Potential for measurement error&lt;br&gt;o Potential for data analysis errors&lt;br&gt;o Logistical challenges if respondents are limited to actual customers</td>
</tr>
<tr>
<td>Survey Experiments &amp; Usability Experiments (with a control condition)</td>
<td>o Confusion created by the firm (and from which the firm typically profits)</td>
<td>o Potential to isolate causation&lt;br&gt;o If subjects are not limited to actual customers, feasibility</td>
<td>o Looming impracticability&lt;br&gt;o Weak ecological validity and looming ecological validity challenges&lt;br&gt;o Population validity concerns and looming challenges&lt;br&gt;o Weak external validity due to subject disinterest&lt;br&gt;o Potential for weak internal validity (control stimulus construction challenges)&lt;br&gt;o Potential for measurement error&lt;br&gt;o Potential for data analysis errors</td>
</tr>
<tr>
<td>Copy Tests &amp; Usability Tests (without a control condition)</td>
<td>o Confusion created by the firm (and from which the firm typically profits)</td>
<td>o If participants are not limited to actual customers, feasibility</td>
<td>o Looming impracticability&lt;br&gt;o Weak ecological validity and looming ecological validity challenges&lt;br&gt;o Population validity concerns and looming challenges&lt;br&gt;o Weak external validity due to subject disinterest&lt;br&gt;o Potential for measurement error&lt;br&gt;o Potential for data analysis errors</td>
</tr>
<tr>
<td>Expert Facial Analyses &amp; Usability Inspections</td>
<td>o Confusion created by the firm (and from which the firm typically profits)</td>
<td>o Economical</td>
<td>o Weak population validity&lt;br&gt;o Weak ecological validity&lt;br&gt;o Looming impracticability</td>
</tr>
</tbody>
</table>

1. **Consumer and Customer Surveys (without a Stimulus)**

**Key Concepts:** Surveys of a firm’s actual customers can have exceptionally strong population validity. These surveys can help assess whether a firm has made unfair or abusive use of existing consumer confusion or has failed to obtain informed consumer consent. Surveys of a firm’s actual customers can also have reasonable construct validity for claims regarding confusion created by the firm. Challenging logistical issues to consider include—how to reach a firm’s actual customers, and how to do so soon enough after the relevant transactions so as to minimize the risk that memory lapses could bias survey results.

Consumer and customer surveys are used to ascertain respondents’ existing knowledge, beliefs and perceptions when these are relevant to a legal claim. No “control group” is tested, because no experiment is being performed. However, “control questions”—such as questions designed to identify respondents who are
guessing, yea-saying, confused about what the question is asking, or not paying sufficient attention to the questions—can be included in the survey instrument. Surveys were once disfavored due to their expense, but online modalities have dramatically reduced that expense.\textsuperscript{224}

To be admissible, consumer and customer surveys (as well as survey and usability experiments and copy and usability tests) generally must be conducted according to accepted current social science practices. The party offering a survey should establish that it was conducted in accordance with accepted principles of survey research, e.g., that:

\begin{itemize}
  \item a proper universe\textsuperscript{225} was examined and respondents constituted a representative sample of that universe;
  \item the questionnaire was properly designed to ask clear, non-leading questions, and was fielded in a manner that did not introduce bias or error; and
  \item the data gathered was completely and accurately coded and reported and analyzed correctly.\textsuperscript{226}
\end{itemize}

However, even if there are deficiencies in one or more of these principles, unless the survey is deeply flawed, courts will typically admit it into evidence but discount its weight. Courts have recognized that “[i]t is notoriously easy for one expert to appear to tear apart the methodology of a survey taken by another,” and therefore admonished that “one must keep in mind that there is no such thing as a ‘perfect’ survey.”\textsuperscript{227} Criticism of one expert’s survey by another expert is particularly undermined where the critic did not perform a survey of his or her own.\textsuperscript{228}

\begin{itemize}
  \item \textit{Construct Validity}
\end{itemize}

Customer surveys have strong construct validity for unfair and abusive practices and informed consent legal claims. In many contexts, customer surveys can also have reasonable construct validity for legal claims that require the firm to be the source of customer confusion.

\begin{itemize}
  \item Surveys to ascertain the knowledge or beliefs of actual customers have been used in unfair practices cases to determine whether customers were so confused that they did not realize they had agreed to a transaction at all.\textsuperscript{229} In these cases, the source of the confusion is irrelevant; the legally-relevant construct is taking advantage of consumer confusion.
\end{itemize}

\textsuperscript{224} See, e.g., Joel R. Evans & Anil Mathur, \textit{The Value of Online Surveys}, 15 Internet Research (2006); Krin Irvine et al., \textit{Law and Psychology Grows Up, Goes Online, and Replicates}, 15 Journal of Empirical Legal Studies 320 (forthcoming 2018) (describing costs of online surveys: “Per subject fees on MTurk range from a few cents to a few dollars…. [C]osts for other convenience samples (like Qualtrics) can be in the $5.00 range, and the costs for curated nationally representative samples are in the $10-$20 per complete[d survey] range”); \textit{Overview of Pricing}, https://support.google.com/360suite/surveys/answer/2447244?hl=en&ref_topic=7171859.

\textsuperscript{225} Some social science disciplines refer to this as the “target population.”


\textsuperscript{228} See, e.g., The President and Trustees of Colby College v. Colby College—New Hampshire, 508 F.2d 804 (1st Cir. 1975) (reversing rejection of survey results based on expert criticism when critic had done no survey of his own).

\textsuperscript{229} See, e.g., State v. Vertrue, Inc., 2010 WL 1652448 (customer survey evidence admissible to show that the “vast majority of consumers were billed for memberships that they did not knowingly enroll in, and were charged for memberships that they never wanted or used”); Federal Trade Commission v. Inc21.com Corp., 745 F. Supp. 2d 975, 1001 (N.D. Cal. 2010), affirmed, 475 F. App’x 106 (9th Cir. 2012) (finding that defendants engaged in unfair and deceptive practices where survey of defendants’ “customers” found that 97% did not believe they had bought the service for which they were being charged and 95% did not even realize they were being charged, presumably because the charge was buried in their telephone bills).
Surveys of a firm’s actual customers have also been used to assess whether a firm has created consumer confusion. For example, a survey of a firm’s customers can be used to demonstrate that the firm is creating false impressions when those impressions are the product of unrecorded sales pitches, such that the precise words used are unknown and cannot be subjected to a survey experiment or copy test. In such cases the court makes the reasonable assumption that customers who bought the product or service are likely to have developed beliefs about its costs and benefits from the sales pitches the customers received.

A similar use of customer surveys can be found in antitrust cases to demonstrate customer confusion that prevents market competition. To constitute an antitrust violation, the plaintiffs in these cases are required to demonstrate that customers are confused, that the firm’s communications led to their confusion, and that this confusion effectively prevented competitors from entering the market. Thus, the relevant construct for this type of case is customer confusion caused by the firm. The usual method for demonstrating causation is an experiment. However, as one court explained, a survey of customer knowledge and beliefs is sufficient to demonstrate causation when there is no reason to believe customers formed their inaccurate beliefs based on anything other than the firm’s communications with the customers:

[A] reasonable jury could reach the following three conclusions. First, [the defendant] used misleading and coercive marketing and sales techniques when presenting [the service] to customers. Second, there is widespread confusion among [service] subscribers in Florida concerning the existence of [the service], the need for [the service], and the scope of coverage provided by [the service]. Third, Florida consumers received most, if not all of their information concerning [the service] from [the defendant]. The third conclusion is critical as it minimizes the probability of an alternative cause. Without evidence of an alternative cause, a reasonable jury could conclude that [the defendant’s] marketing and sales techniques caused consumer confusion.

A final example comes from the remedial stage of a case in which the court had found that the defendant engaged in a deceptive advertising campaign that had confused consumers into believing that its product had attributes it did not have. As part of the remedial decree, the defendant was required to hire an expert to perform a series of surveys to assess whether consumers remained confused about the product’s attributes after the deceptive marketing had ended. The defendant was
required to engage in corrective labeling/advertising until consumer confusion had substantially dissipated.\textsuperscript{235}

Looking to the future, as firms disseminate hundreds or even thousands of different micro-targeted websites and advertising, customer surveys may become a particularly desirable and practicable approach to assess customer confusion. As with unrecorded telephone calls, all micro-targeted communications cannot realistically be reproduced and subjected to survey or usability experiments. Yet, if a firm’s customers mistakenly believe prices are lower or benefits are higher than is true for the firm’s product or service, the most plausible explanation will often be that this confusion was caused by the firm. Competitors would generally have an incentive to lead consumers to believe that the opposite—that the firm’s offerings are more costly or less beneficial than they truly are. Other parties would generally be agnostic about customer understanding of key product or service facts. Customer confusion favoring a firm’s product or service is thus in many situations likely to have arisen from customers’ direct exposure to firm communications or from customers’ interactions with other consumers who formed their beliefs based on firm communications.

- **Strengths**

**Potential for Strong Population Validity.** Consumer and customer surveys can achieve a degree of population validity beyond what any conventional methodologies for assessing consumer or customer confusion can produce. The survey respondents (as well as respondents in survey or usability experiments and copy or usability tests) are likely to be representative of the relevant population\textsuperscript{236} provided that the survey employs the following:

1. the appropriate population universe;\textsuperscript{237}
2. probability sampling of a sample frame that does not differ from that universe in any relevant respect (thus avoiding coverage error);\textsuperscript{238} and
3. no relevant differences between those who do and do not complete the survey that would result in a non-response bias.\textsuperscript{239}

\textsuperscript{235} In trademark cases in the U.S., surveys regarding consumer beliefs are used for two purposes. First, to assess genericness—if consumers perceive a mark to be generic rather than associated with a particular brand the mark cannot receive (or can no longer receive) trademark protection. Second, to assess secondary meaning—a particular mark that is insufficiently distinctive to be registered as a trademark can, over time, become so strongly associated in consumers’ minds with a particular brand that it acquires trademark protection. The genesis of consumers’ perceptions about whether a mark is generic or has acquired a secondary meaning is not legally relevant. See, e.g., Herman Miller, Inc. v. Palazzetti Imports & Exports, Inc., 270 F.3d 298, 312 (6th Cir. 2001) (“Because the determination of whether a mark has acquired secondary meaning is primarily an empirical inquiry, survey evidence is the most direct and persuasive evidence.”).

\textsuperscript{236} The researcher can take the results of such a survey and, based on the confidence level desired (usually 95%), the size of the sample, and the variability in the response being measured, report a confidence interval for the value in question. Shari Seidman Diamond, *Reference Guide on Survey Research*, in *Reference Manual on Scientific Evidence* 359, 381 (Federal Judicial Center: 3rd ed. 2011).

\textsuperscript{237} Note that screening questions at the start of the questionnaire are used to ensure that respondents are members of the population of interest and to eliminate respondents who might have specialized knowledge of the topic, and therefore would not represent an average consumer.

\textsuperscript{238} Courts sometimes find nonprobability (convenience) sampling to be acceptable. See, e.g., National Football League Props., Inc. v. New Jersey Giants, Inc., 637 F. Supp. 507, 515 (D.N.J. 1986) (admitting survey evidence taken from a nonprobability sample on the grounds that “results of these studies are used by major American companies in making decisions of considerable consequence”). However, a survey that uses probability sampling would generally be entitled to more weight.

\textsuperscript{239} Sometimes it is possible to statistically compensate for uneven response rates between different subgroups in the population universe, by assigning different weights to responses from different groups. Shari Seidman Diamond, *Reference Guide on Survey Research*, in *Reference Manual on Scientific Evidence* 359, 384 (Federal Judicial Center: 3rd ed. 2011). Note also that high payoff low probability lotteries have sometimes been found to increase response rates significantly. See, e.g., Henry Sauermann & Michael Roach.
A survey of a firm’s actual customers by definition samples the appropriate population universe. The sampling frame might be that entire universe of actual customers, although logistical constraints might require the use of a narrower frame.

At one time surveys conducted online had weak population validity. However, as of 2015, nearly 90% of U.S. adults spend time online.\textsuperscript{240} Survey work utilizing online panels are now prevalent in litigation as well as in academic research. In a recent case, the Federal Trade Commission compared the use of random-digit-dialed telephone surveys and an online methodology (“Google Survey,” a relatively recent innovation by Google) for achieving population validity.\textsuperscript{241} The Commission opined that the set of internet users who serve as the sampling frame for Google surveys is an appropriate sample frame for the relevant population, which was all adult U.S. consumers. The low per-respondent expense of the online study also facilitated the use of micro-survey techniques (asking each respondent a single question). Micro-surveys can reduce subject fatigue, which might otherwise reduce study completion rates and potentially introduce non-response bias.

On the other hand, some firms target unfair, deceptive, or abusive practices at older, minority, less-educated, rural, and/or and poorer Americans, all of whom are disproportionately offline.\textsuperscript{242} For these cases, surveys might use telephone, mail, or in-person interview modalities to boost population validity. Of course, no matter the subpopulation targeted, if the firm has been transacting with consumers online, the relevant population universe will probably be online.

- **Weaknesses**

  **Potential for Measurement Error.** Consumer and customer surveys are susceptible to measurement error, although errors can often be minimized through a carefully crafted and well-fielded survey instrument (questionnaire).

  Unclear or leading questions can seriously undermine the results of consumer and customer surveys, survey and usability experiments, and copy and usability tests.

  - For example, in one case a consumer debtor claimed that a debt collector’s form dunning letter, as an unfair pressure tactic, misleadingly suggested that a debtor had to “act now” or would not be able to settle the matter. The appellate court in that case ruled inadmissible survey evidence that the plaintiff attempted to use to demonstrate that a significant portion of debtors would have believed the letter to convey this false message. The court explained that the survey improperly employed leading questions.\textsuperscript{243}

An interviewer whose body language or facial expression conveys acceptability of some responses and not others can have an equally damning effect.


\textsuperscript{243} Evory v. RJM Acquisitions Funding, L.L.C., 505 F.3d 769, 778 (7th Cir. 2007). See also, e.g., Valador, Inc. v. HTC Corp., 242 F. Supp. 3d 448, 465 (E.D. Va.), affirmed, 707 F. App’x 138 (4th Cir. 2017) (excluding results from a survey experiment for several reasons, including the use of suggestive or leading questions).
Acquiescence bias or “yea-saying,” a tendency to respond in the affirmative to survey items irrespective of the content of the question, can bias results unidirectionally.244 Relatedly, if subjects believe that particular responses are more socially acceptable, social desirability bias could skew results towards more socially desirable answers.245 Order effects (effects caused by the order of the questions246 or of the responses247) can be a source of measurement error.

Instrument design, modality, presentation, and timing can make a difference here.248 For example, best practices involve pre-testing surveys, meaning testing the survey with a small number of respondents prior to fielding the final version of the survey.249 By debriefing respondents, researchers can identify instructions, questions, and answer choices that are unclear or ambiguous and can uncover issues with the flow and length of the instrument. Possible nonresponse problems can also be explored. It may then be possible to address these issues before fielding the survey to the entire sample. In the finalized questionnaire, control questions can test again whether respondents are interpreting the questions and the answer choices in the way in which the researcher intends them to be interpreted.250

Control questions can also help identify yea-sayers and varied question framing—asking some respondents a question in the positive and others in the negative—can cancel out some yea-saying effects. Question topic order and (for closed-ended questions) response order can be rotated to reduce order effects.251 Online micro-surveys, which ask only a single substantive question, avoid order effects, provided that non-substantive questions in these surveys do not produce order effects.252 Interviewers can be kept blind to the purpose of an in-person survey to mitigate the effects of interviewer bias, or online modes of data collection can be used to...
eliminate interviewer bias and mistakes. Online surveys can also be designed to make respondents more comfortable admitting when they do not know an answer, whereas respondents may feel social desirability bias to guess when a live interviewer asks them a question to which they think they should know the answer.

A challenging issue for measuring customer or consumer knowledge of or confusion about key product and service facts is question design. Both open-ended and closed-ended questions are often used to assess consumer knowledge, beliefs, and level of confusion. Both formats have strengths as well as weaknesses, and they often do not produce the same results.

Open-ended questions are good at eliciting salient and top-of-mind responses and can avoid the bias that can be introduced by closed-end question response choices. However, open-ended questions may fail to elicit respondent knowledge that is not top-of-mind but that would have come to mind with prompting. They can also be difficult to code because respondents intending to convey the same idea may use different words to express themselves.

Clearly worded and well formulated closed-ended questions are good at prompting respondents to focus on a larger range of response options rather than just those that are top-of-mind. They also lend themselves more easily to coding and quantitative analyses. However, the specific language in the response choices may lead respondents and introduce bias. Further, they may produce responses that are merely guesses, rather than reflecting the respondent’s actual knowledge. For example, a closed-ended question could ask respondents to identify a service’s benefits. The presence of answer choices could lead respondents to guess about potential benefits, but an answer to the same question asked in an open-end format could fail to elicit a benefit that the respondent believes the service has but could not recall without the reminder of a closed-end response choice.

Researchers often use both open-ended and closed-ended questions to take advantage of their offsetting strengths and weaknesses. Open-ended questions about a topic are typically asked first, followed by more focused closed-ended questions on that topic. This “funnel” approach is designed to ensure that closed-ended questions do not bias answers to open-ended questions, and is viewed favorably by judges.

Another method to reduce guessing is to ask a filter question before asking a substantive question. If a participant states that he or she believes the product has a particular cost or benefit, follow-up questions can probe his or her impression of the parameters of that cost or benefit. For example, “Do you know whether you will be charged a fee if you overdraw your account?” could be asked before asking about the amount of the fee and conditions under which it is imposed. If a respondent states that a loan comes with “insurance,” follow-

253 See, e.g., Jeff Miller, Online Marketing Research, in The Handbook of Marketing Research 110-131 (R. Grover and M. Vriens, eds., Sage Publications: 2006) (asserting that online surveys are often preferable to in-person surveys because they elicit more truthful answers (due to anonymity) and eliminate any bias that an interviewer could introduce).

254 Neil Malhotra et al., The Relationship between Nonresponse Strategies and Measurement Error: Comparing Online Panel Surveys to Consumer Belief Surveys, in Online Panel Research: A Data Quality Perspective (2014) (finding that internet survey respondents are more likely to admit that they do not know an answer, and telephone survey respondents are more likely to pick the middle option, presumably to appear competent).

255 Howard Schuman & Stanley Presser, Questions and Answers in Attitude Surveys (Sage Publications: 1996).

256 See Jon A. Krosnick & Stanley Presser, Questionnaire design, in Handbook of Survey Research 263, 266-68 (J. D. Wright & P. V. Marsden, eds., Emerald Group: 2nd ed. 2010).


258 See, e.g., In the Matter of Telebrands Corp., 140 F.T.C. 278, 316-17 & n.36 (2005) (approving of the funnel approach). See also Consumer Financial Protection Bureau, Arbitration Study: Report to Congress, pursuant to Dodd–Frank Wall Street Reform and Consumer Protection Act § 1028(a), § 3.4 (March 2015) (using open-ended questions followed by closed-ended questions to assess consumer confusion about arbitration clauses in their credit card contracts).
up questions can assess what the respondent believes that insurance covers and does not cover. In addition, a “do not know” answer can be provided as a quasi-filter, although sometimes respondents are reluctant to admit they do not know an answer to a question, particularly when they think they should know the answer.

Many sources of measurement error can be made worse when respondents do not devote enough cognitive effort to the task. Respondents know that their answers are unlikely to ultimately affect their own lives. Respondents may therefore be tempted to answer survey questions too quickly, without taking the cognitive retrieval and processing time needed to ensure accuracy, a phenomenon called “satisficing.”

However, many studies show that satisficing is less of a problem than one might think. Even online respondent pools that complete many surveys appear to do so carefully, and the level of compensation need not be high to generate significant respondent effort.259 For confusion audits of actual customers, the questions concern transactions in which respondents have engaged, meaning that the questions inherently have personal relevance, a factor that has been found to reduce satisficing.260 Sometimes satisficing is caused by respondent fatigue; keeping survey length modest and ensuring that questions and closed-end answers are short and easy-to-follow can be useful when respondent fatigue is a problem.261

In situations where a live interviewer might lead some respondents to feel social pressure to respond quickly, online modalities might reduce that feeling. To physically prevent online respondents from moving through questions too quickly, an online survey instrument can use programmable timers to prevent the respondent from reaching the next question until a pre-determined amount of time has been spent on a question.262 Manipulation checks such as control questions can also be used to screen out satisficers.263

**Potential for Data Analysis Errors.** Beyond issues of population validity and measurement error, methodological defects at the data analysis stage (e.g., poor coding of answers to open-ended questions or errors in statistical analyses of results), can also impair the validity of survey evidence.264 In litigation in the U.S., the adversary system aims to protect against poor data analysis through transparency; by making raw survey data available to the opposing party,265 the opponent can employ experts to check for coding errors and perform their own statistical re-analyses.

**Logistical Challenges for Surveys of Actual Customers.** Surveys of a firm’s actual customers present at least two logistical challenges: locating the customers and doing so soon enough to obtain unbiased responses.

---


264 With live interviewers, there can also be quality control issues, in that interviewers may be tempted to fake responses so as to reduce their workload. Audits are often used to try to eliminate faked responses, but a cleaner way to deal with the problem is to eliminate live interviewers.

265 This data is not always released to the public.
First, it may be difficult to locate the customers who have engaged in the relevant transaction with the firm. If the transaction is engaged in at high enough rates throughout the population, and the firm has a large enough market share, it is possible that a researcher could survey a large representative sample of the adult population and use screening questions to narrow respondents to those who engaged in the particular transaction with the particular firm.266 With a smaller relevant population, this method could quickly become quite inefficient.

For consumer financial services transactions, it is likely that firms retain records that are sufficient to identify and locate their customers. In litigation in the U.S., firms can be required to confidentially provide plaintiffs’ experts with all relevant customer records or with a representative sample of these records selected using a method agreed upon by the parties or ordered by the court. Outside of litigation, a regulator would need the power to obtain customer records (or representative samples of those records) from the firm.

A second challenge is posed by the passage of time between the time customers engage in transactions and the time they answer the survey questions. The legally relevant confusion for most types of claims is confusion at the time the consumer makes the decision to buy the product or service or to take another action from which the firm profits. A confusion audit seeks to discover whether customers knew the relevant key fact about the product or service when they engaged in decision-making about the transaction.267

On the one hand, some respondents might have learned more about the relevant key facts in this time period, whether through experience or otherwise, and consumers probably cannot determine whether their beliefs about the transaction were formed before or after engaging in it through introspection.268 Knowledge acquired after the transaction and before answering the survey would tend to bias confusion rates downwards.

On the other hand, some respondents might have forgotten some key facts between the time of the transaction and the time of the survey, potentially biasing confusion estimates upwards.269 For financial products and services, there is some evidence that consumers who do understand key facts at the time of decision do not always strongly encode these facts into long-term memory.270 However, it is not clear that respondents who are asked about a key fact sometime after a transaction and who do not at the time of the survey know that fact can accurately assess whether (a) they never knew that fact or (b) they knew it at the

---

266 Researchers in the U.S. have done this by asking consumers to identify one of their credit card issuers and their beliefs as to whether their credit card agreement with that issuer contains arbitration clauses or class action waivers, and then comparing this to hand-collected credit card agreements from the major issuers. Jeff Sovern et al., “Whimsy Little Contracts” with Unexpected Consequences: An Empirical Analysis of Consumer Understanding of Arbitration Agreements, 75 Maryland Law Review 1 (2015-2016).

267 A policymaker seeking to ensure that switching costs are low would also want consumers to know the price they are paying for financial services on an ongoing basis. U.S. consumer protection and fair competition law does not currently concern itself with this.

268 See, e.g., Richard E. Nisbett & Timothy D. Wilson, Telling More Than We Can Know: Verbal Reports on Mental Processes, 84 Psychological Review 231 (1977); D.J. Herrmann, Know Thy Memory: The Use of Questionnaires to Assess and Study Memory, 92 Psychological Bulletin 434 (1982), http://dx.doi.org/10.1037/0033-2909.92.2.434 (reviewing studies finding that people’s beliefs about their memory performance are often inaccurate).

269 For an overview of memory issues in survey research, see Roger Tourangeau, Remembering What Happened: Memory Errors and Survey Reports, in The Science of Self-Report: Implications for Research and Practice 29 (Arthur A. Stone et al., eds., Psychology Press: 1999). Memory problems have been documented with respect to the prices of retail goods, in that sometimes survey respondents who knew prices sufficiently well to comparison price shop at the retail store are unable to remember those prices with much precision when they respond to surveys, even surveys conducted soon after the transaction. See, e.g., Kent B. Monroe & Angela Y. Lee, Remembering Versus Knowing: Issues in Buyers’ Processing of Price Information, 27 Journal of the Academy of Marketing Science 207 (1999).

time they engaged in the transactions but cannot now recall it. If a customer asked “Did you buy credit insurance with your car loan?” answers “Do not know,” this answer is ambiguous. It could mean that the customer does not recall whether he or she bought credit insurance or that at the time he or she obtained the loan he or she did not know whether the transaction included the purchase of credit insurance. Adding “I knew when I bought it, but I no longer recall” as an answer choice will not help if respondent introspection cannot distinguish between the two mental states.

There are a number of ways memory problems can potentially be mitigated, but doing so could introduce other biases. For example, actual customers could be given access to materials they encountered during the purchase process to use as a memory aid during the survey. This will not directly assist with recall of, for example, information the customer received orally from a salesperson. However, memory often improves when people are provided with even unrelated contextual clues that help people mentally recreate the relevant personal experience.

One challenge for this approach is determining what materials the customer encountered during the transaction decision process. As explained in connection with unaided judicial judgment, in the twenty-first century firms are increasingly using a multitude of micro-targeted communications. Even if a firm’s disclosure document remains unchanged across customers for a period of time, the firm communications surrounding that disclosure may change. For memory aid purposes during a survey, the materials probably do not need to be exactly the same as those used by the firm in the field. However, to the extent that respondents use the materials as more than a memory aid, and instead form new beliefs based on the materials, the survey will in effect operate as a copy test rather than a survey of customers’ existing knowledge or beliefs. This could lead customers to report as beliefs things they did not actually understand at the time of the transaction, which could bias results. To obtain the precise materials the customer used in the course of his or her actual transaction so as to avoid this bias would require, at a minimum, significant cooperation from the firm whose customers are being tested.

Another potential way to minimize memory problems would be to survey customers as to their knowledge and perceptions close in time to the transaction at issue. How soon after a transaction a survey would need to take place might depend on how persistent consumer knowledge of key facts of that transaction must be for the consumer to use that knowledge to make informed autonomous decisions that could enhance substantive competition, as discussed above in connection with confusion constructs. For complex decisions that require consumers to gather and process information and reflect upon tradeoffs over time, knowledge of key facts probably needs to persist in memory for that entire period, meaning the knowledge would probably be somewhat strongly encoded into long term memory and retrievable for some time. For simpler decisions, the information must be encoded to be used, but does not need to persist to make good decisions. It is therefore possible that to avoid bias due to memory errors, customer surveys would need to be performed very soon after simple transactions but could occur sometime later for complex transactions.

What might be most helpful would be for financial firms to facilitate access to their customers in real time so that independent experts can perform the surveys immediately after the pertinent transactions.

---


Alternatively, firms might be required to survey representative samples of their own customers at the time those customers engage in the relevant transactions, using the firms’ own data and processes for reaching customers. This would require tight oversight by a regulator, perhaps through independent experts selected or approved by the regulator or experts in survey design and implementation employed by the regulator, working on the ground alongside firm employees helping to field the surveys.

The ease with which customers can be accessed soon after engaging in a transaction could depend on, e.g., survey modality and respondent communication habits. When conducted online, and in most instances when conducted by telephone, a survey could be presented to customers on the heels of a transaction (again, with substantial cooperation from the firm). Thus, respondents could answer survey questions very soon after engaging in the pertinent transaction. If in-person interviews are needed to draw respondents engaged in offline transactions, a researcher in theory could camp out next to a firm’s retail location and ask customers questions on their way out the door. However, when the firm realizes this is happening, the firm is likely to spend (more) effort to be certain that customers understand the key facts of their transactions with the firm, undermining the ecological validity of the survey.

* * * * *

Despite these potential issues, consumer surveys, and even more so customer surveys, can overcome population validity issues that conventional forms of consumer perception evidence cannot. For cases in which customer confusion based on the entirety of what each customer believes about the transaction is the legally-relevant construct, a customer survey could be optimal. As business communications become increasingly personalized, rendering other methodologies impracticable, customer surveys may also be the best way to capture confusion caused by a particular firm’s communications.

Logistics are not easy. In particular, surveying a firm’s actual customers close in time to when the customer engaged in the transaction could prove to be important and could require substantial cooperation from firms. However, ecological validity is not a problem for these surveys because customers are not asked to assess an experimental manipulation but rather to answer questions in the context of their own lives. The potential for measurement errors is always present but can be minimized with careful instrument design and fielding. Data analysis errors can be corrected.

In sum, properly done, customer surveys have the potential to provide solid evidence of the presence or absence of confusion among a firm’s customers about the customers’ transactions with the firm.

2. **Survey Experiments and Usability Experiments (with a Control Condition)**

**KEY CONCEPTS:** Survey and usability experiments can provide evidence that a firm’s specific practices are likely to cause consumer confusion. However, twenty-first century real-time micro-targeting of marketing and online interfaces is likely to undermine the ecological and population validity of these experiments for demonstrating this causation. Further, constructing an adequate control can be challenging. Beyond demonstrating causation, this type of research might productively be used to isolate confusing features of advertisements, websites, pricing structures, and product and service designs, so as to guide firms and regulators in their efforts to improve consumer understanding of their transactions.

To be admissible, survey and usability experiments must be conducted according to generally accepted current social science practices.\(^{275}\) This means they must meet the requirements for consumer and customer

---

\(^{275}\) A/B testing is a form of usability experiment practiced throughout industry. For an overview of large-scale A/B testing that can produce statistically significant results, see Alex Deng et al., *A/B Testing at Scale: Accelerating Software Innovation*, in Proceedings of
surveys discussed above (appropriate universe, non-leading questions, careful coding, etc.). In addition, subjects must be randomized into treatment and control groups, and a control must be created to act as a stimulus for the control condition. As with all forms of evidence, courts do not demand perfection of survey or usability experiments.

To perform a survey or usability experiment, members of a treatment group are presented with particular materials that are alleged to be misleading and asked about their impressions of this test stimulus. Members of a control group are asked similar kinds of questions about a control stimulus (typically, the control is the same material shown to the treatment group, with the aspects that are suspected of being misleading eliminated) under similar conditions. One function of the control group is to isolate the causal effect of the test stimulus on consumer confusion, separately from pre-existing consumer beliefs not created by the test stimulus. Any differences between the answers given by the treatment group and the control group are generally ascribed to the test stimulus.

- **Construct Validity**

  The use of a control condition to net out any pre-existing consumer confusion in survey and usability experiments means that the results of these experiments are relevant to legal claims that do not cover taking advantage of pre-existing consumer confusion. For example, in cases alleging deceptive practices or false competitor advertising, these experiments are appropriate for demonstrating what message or information is conveyed to consumers by a specific advertisement, website, or disclosure document. When particular disclosures are required by law to be clear and conspicuous, these experiments could also be appropriate to assess whether the firm’s presentation of those disclosures meets the legal requirements. However, when the relevant consumer confusion construct includes confusion from which the firm profits regardless of the source of that confusion, such as claims for unfair or abusive conduct, these experiments can sometimes net out the very background confusion of which firms are taking unfair or abusive advantage. On the other hand, where the court concludes that consumers are unlikely to have brought pre-existing confusion to their interactions with the firm, such that any measured confusion is likely attributable to the firm, survey and usability experiments might be able to test for that confusion.

- **Strengths**

the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (Association for Computing Machinery 2007), https://doi.org/10.1145/3077136.3082060.

276 Control conditions are also used to screen out the potentially biasing effects of survey artifacts; because both groups are asked the same questions, if questions are leading or ambiguous, or respondents are inattentive or guessing, responses from both groups should be affected equally. However, because a control condition will also screen out pre-existing confusion from which the firm may be profiting, controlled experiments can be problematic in cases in which that type of mis-selling is the relevant legal construct. Other techniques, such as control questions and filter questions, can be used to minimize the potentially biasing effects of survey artifacts instead of a control condition.

277 See, e.g., Southland Sod Farms v. Stover Seed Co., 108 F.3d 1134, 1142–43 (9th Cir. 1997); American Home Products Corp. v. Johnson & Johnson, 577 F.2d 160 (2nd Cir. 1978); Rexall Sundown, Inc. v. Perrigo Co., 651 F. Supp. 2d 9 (E.D.N.Y. 2009); Mutual Pharmaceutical Co. v. Iax Pharmas. Inc., 459 F. Supp. 2d 925 (C.D. Cal. 2006); Novartis Consumer Health v. Johnson & Johnson-Merck Consumer Pharmaceuticals, 129 F. Supp. 2d 351 (D.N.J. 2000). Survey experiments also test for the right confusion construct in trademark infringement actions. Absent evidence of a firm’s intent to infringe on another’s trademark, survey experiments are required for the trademark holder to prove that the non-holder’s marketing has confused consumers into thinking the non-holder is the holder or is associated with the holder. See, e.g., Union Carbide v. Ever-Ready, Inc., 521 F.2d 366 (7th Cir. 1976); Squirtco v. Seven-Up Co., 628 F.2d 1086 (8th Cir. 1980).

278 See, e.g., Expert Report of Tülin Erdem, Ph.D., Federal Trade Commission v. DirecTV, Case 4:15-cv-01129-HSG, Document 261-3 (Filed 01/03/17) (reporting results of a survey experiment performed on one of the defendant’s print advertisements and results of a usability experiment performed on one of defendant’s websites).
Potential to Isolate Causation. Survey and usability experiments are forms of randomized, controlled trials, which are generally considered the “gold standard” for establishing a causal relationship between a treatment and a measured outcome. When the construct relevant to the litigation is the effect on consumers of a particular advertisement, website, or other communication used by a firm, survey and usability experiments have the potential to directly demonstrate that causation. If the control stimulus is properly constructed, background beliefs with which subjects approach the transaction can be measured so as to isolate the causal effect of the stimulus on subject beliefs.

Feasibility. Subjects in survey and usability experiments are usually taken from the population of consumers who are potential users of the product or service in question. Therefore, subjects are less challenging to locate than if the relevant population were limited to actual customers who purchased a particular product or service from a particular firm.

- Weaknesses

Looming Impracticability. For the same reasons discussed above in conjunction with unaided judicial judgment, the proliferation of micro-targeted quickly-changing business-to-consumer communications may render survey and usability experiments impracticable for litigation purposes. Any methodology that attempts to measure consumer confusion by looking at exemplars of a firm’s communications faces the same challenge. In instances when hundreds or even thousands of micro-targeted versions have been deployed, experiments testing even a fraction of these communications would be burdensome. Testing of representative samples of communications would mitigate this burden, but the sample size needed to produce meaningful results could still be onerous.

Weak Ecological Validity and Looming Ecological Validity Challenges. Ecological validity is a significant concern for survey and usability experiments. Although many survey experiments were once conducted at shopping malls, in the physical vicinity of the place where a consumer might purchase the product, and some are performed using simulated versions of store shelves where the product might be sold, these are still artificial environments that are stripped of much of the context that affects consumers in the real world.

A control condition can eliminate some of the effects of the change in context, because both the treatment and control groups are exposed to an artificial environment. But some contextual features that can be very powerful are necessarily absent from the experiment. Frequently, particularly when salespeople have an incentive to make sales, salesperson activity is likely to be an overwhelmingly influential aspect of the real-world environment, yet is too variable to replicate accurately in an experiment. In addition, consumer perceptions are influenced by subtle aspects of an environment, such as physical placement or timing of receipt of a communication, and these subtle aspects are difficult to identify and replicate.

Website usability testing can be performed more naturalistically, because typically no live salesperson is present when the customer is online. Yet even a usability experiment misses some real world context. A website can be precisely replicated, but the process by which the consumer reaches it cannot, and this process could strongly influence the messages the consumer takes from the website. One court discounted the results of a survey that attempted to measure whether consumers would be misled by products sold on a particular website in part because the normal process of reaching the website was not replicated, explaining:

[P]eople do not come to websites randomly, and they do not move within websites randomly. A great majority of Internet users arrive at a particular website after searching specific terms via an Internet search engine or by following links from another website. The user makes a

See, e.g., Conopco, Inc. v. Cosmair, Inc., 49 F. Supp. 2d 242, 253 (S.D.N.Y.1999) (referring to a survey experiment: “To have substantial probative value, a survey ... must ... be designed to examine the impression presented to the consumer by the accused product. Therefore, a survey must use the proper stimulus, one that tests for confusion by replicating marketplace conditions.”).
judgment based on contextual cues—what is shown about a prospective website from the text of a search result or what is said about a prospective website in the hyperlinked words and surrounding text of the website currently being viewed—in determining where to surf next. He moves from website to website, he moves within websites, and he performs actions such as signing a petition—or buying a product—by making choices based on what he sees and whether what he sees leads him to believe that going to the next page or following a link to another website will bring him to something he is interested in seeing, doing or buying.\textsuperscript{280}

Further, the website simulation itself must be exact; one study found that adding even a 15-second distraction before the display of a disclosure on a website changed subjects’ responses.\textsuperscript{281} Therein lies the rub. Recall that online communications are increasingly micro-targeted to consumers in real time. For example, an online advertisement stressing the purported beautifying effects of a product might be sent to women on Monday mornings, when they feel least attractive, but the same product might be pitched with a health theme in late December and early January, when New Years’ resolutions are at their height. Financial products will similarly find particular consumers more or less receptive at different times and against different economic news and personal events. Matching subject environments to the particular website that would be shown in that environment in the real world is an impracticable if not impossible task for a third-party researcher to perform. Ecological validity is thereby undermined.

The courts have yet to tackle the problem of matching experimental subjects to the micro-targeted materials they would have viewed online, but they have addressed a similar issue in addressing consumer confusion created by unrecorded in-person and telephone interactions, discussed above. With micro-targeting, the entire sales environment is coming to more and more closely resemble the one-on-one sales context.\textsuperscript{282} Just as salesperson behavior cannot be accurately replicated in an experiment, so too the environments consumers are in at the moment they are transacting on a particular firm’s website cannot be accurately replicated. When faced with this conundrum in the context of unrecorded sales pitches, courts have accepted surveys of actual customers rather than survey experiments for demonstrating that firms have created the misimpressions and confusion identified in the surveys.\textsuperscript{283}

\textbf{Population Validity Concerns and Looming Challenges.} Micro-targeting of consumer-facing materials also raises a population validity problem for survey and usability experiments. It may be possible to obtain an acceptably representative sample of all potential consumers in the relevant market and split them between test and control conditions. But the actual population that encounters any particular website or pop-up advertisement in the real world is not all potential consumers in the market for the product or service being sold. Rather, each subpopulation receives its own micro-targeted communication from the firm. Matching subjects’ personal characteristics to the particular website that would be shown to that potential customer is impracticable if not impossible for a third-party to do. As with ecological validity, population validity in survey and usability experiments under these conditions will be weakened.

Even without accounting for twenty-first century trends, population validity will likely be weaker for survey or usability experiments than for actual customer surveys. The population from which subjects are selected for these experiments are potential purchasers of the specific product or service in question, but they might never become actual customers. That someone who never engages in the transaction would have been misled if he


\textsuperscript{281} Idris Adjerid et al., Sleights of Privacy: Framing, Disclosures, and the Limits of Transparency, Symposium on Usable Privacy & Security 1 (2013).

\textsuperscript{282} For further discussion, see Lauren E. Willis, Performance-Based Remedies: Ordering Firms to Eradicate Their Own Fraud, 80 Law & Contemporary Problems 7 (2017).

\textsuperscript{283} See, e.g., Schering Corp. v. Pfizer Inc., 189 F.3d 218 (2nd Cir. 1999), as amended on reh’g (Sept. 29, 1999).
or she had used the website through which the transaction is performed is not itself a problem for consumer autonomy and market competition. More importantly for research purposes is that there could be unobservable differences between those consumers who become actual customers and those consumers who do not, and these unobservable differences could affect confusion rates.

**Weak External Validity due to Subject Disinterest.** Some have posited that the most important difference between consumer confusion experiments and the real world is that subjects are not actually shopping for themselves, and therefore lack the intrinsic motivation consumers in the real world have to pay attention to the stimulus and engage in cognitive retrieval and processing. One U.S. court, in a case involving a survey experiment, has asserted that all survey results are imperfect because “people are more careful when they are laying out their money than when they are answering questions.” This presents the “satisficing” issue discussed above.

Certainly, subjects react differently to experiments than they do to real-life stimuli. But some subjects might actually be more likely to attend to the stimulus and give it more conscious thought in an experiment than they would in the hurly-burly flow of life. Assessments of the use of survey experiments for academic purposes have therefore concluded:

[E]xperiments are normally carried out in such a way that virtually everyone receives the message. The typical experiment thereby avoids a major obstacle standing in the way of communication effects, namely, an inattentive audience, lost in the affairs of private life. By ensuring that [stimuli] reach their intended audiences, experiments may exaggerate their power.285

Courts have noticed the same phenomenon, i.e., that subjects sometimes pay more attention to the allegedly deceptive materials during the experiment than they would in real life.286

What is the effect of imperfect ecological validity on the results of experiments that test consumers’ perceptions of, and confusion about, key product and service facts? If subjects pay more attention to stimuli than actual consumers, experimental results may overstate consumer understanding and understate consumer confusion. For example, marketing materials produced by firms often contain fine print disclosures in an attempt to avoid liability for deceptive practices. Without distractions, subjects are probably more likely to read and attend to these during the experiment than consumers do in the real world, such that confusion rates will be biased downwards. Even for information that is in large print, fewer distractions can lead to better subject understanding. As noted in the introduction to this Report, one study found that adding even neutral experimenter banter was enough to prevent subjects from using a disclosure to understand key facts about a mortgage loan.

284 Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club Ltd. Partnership, 34 F.3d 410, 416 (7th Cir. 1994).
285 Donald Kinder, *Curmudgeonly Advice*, 57 Journal of Communication 155, 157 (2007). Empirical evidence supports this conclusion. See Jason Barabas & Jennifer Jerit, *Are Survey Experiments Externally Valid?* 104 American Political Science Review 226 (May 2010) (comparing survey experiments to observational field studies and finding that subjects in the experiments appear to have paid more attention to the stimuli and processed it more deeply than people do when encountering comparable stimuli in the real world, and suggesting that “multiple stimuli (competing messages) in the real world ... reduce the effect of any given message”).
286 See, e.g., ConAgra, Inc. v. George A. Hormel & Co., 784 F. Supp. 700, 734 (D. Neb. 1992), affirmed, 990 F.2d 368 (8th Cir. 1993) (finding survey experiment results suspect because subjects spent an average of 8 to 10 minutes examining the product display used for the stimulus and opining: “[C]onsumers do not sit and stare at products, for purposes of answering questions, for substantial periods of time when they are shopping.”).
Potential for Weak Internal Validity (Control Stimulus Construction Challenges). Another challenge for survey and usability experiments lies in the construction of a control stimulus. In business-to-business false advertising cases, the plaintiff competitor can use its own advertising as a control, thereby eliminating any background confusion from which it also profits.288 Even for deceptive practices cases, sometimes all that needs to be done is to remove a few bits of text or graphics. A case from the non-financial consumer product market provides a simple example. Using a survey experiment, researchers determined that removing the true but misleading statements “100% naturally sourced sunscreen ingredients” and “natural protection” from a sunscreen lotion bottle’s label was enough to dramatically reduce consumer confusion. With the statements, about 90% of subjects believed that the entire lotion, not just the active “sunscreen” ingredients, was “100% natural.” Without the statements, fewer than 12% of the subjects believed the lotion was 100% natural.289 Removing just seven words from the label reduced consumer confusion by nearly 80 percentage points.

Where more dramatic changes must be made to eliminate consumer confusion caused by the test stimulus, it can be difficult to create an adequate but proper control. Firms sometimes submit survey experiments in which the primary difference between the actual marketing or other materials and the control stimulus is the addition of a disclaimer. On the surface, this complies with the admonition that “[a] control should not be so different that it artificially lures respondents away from revealing the effect of noise.”290 However, disclaimers have frequently been found to have little impact on consumer beliefs,291 and courts have sometimes rejected control stimuli constructed in this manner.292

A source produced by the U.S. Federal Judicial Center and frequently consulted by jurists is the Guide to Survey Research. This source explains that an appropriate control stimulus should “share[] as many characteristics with the experimental stimulus as possible, with the key exception of the characteristic whose

288 See, e.g., Expert Report of Dr. Charles D. Cowan, Armor Screen Corp. v. Storm Catcher, Inc., 2009 WL 1557680 (S.D. Fla.). Creating controls to test likelihood of confusion in trademark cases is also not too difficult, given that the trademark-protected elements and only the trademark-protected elements must be removed to create the control.


291 See, e.g., Aaron S. Kesselheim et al., Mandatory Disclaimers on Dietary Supplements Do Not Reliably Communicate the Intended Issues, 34 Health Affairs 438 (2015), doi: 10.1377/hlthaff.2014.0515 (conducting a systematic review of all empirical studies of consumer responses to health-related disclaimers and finding that rigorous studies found these disclaimers ineffective); Manoj Hastak & Dennis Murphy, Effects of a Bristol Windows Advertisement with an “Up To” Savings Claim on Consumer Take-Away and Beliefs *4 (May 2012), https://www.ftc.gov/sites/default/files/documents/reports/effects-bristol-windows-advertisement-savings-claim-consumer-take-away-beliefs/120629bristolwindowsreport.pdf (finding that adding a disclaimer acknowledging that consumers on average saved only 25% did not reduce the misimpression created by an advertisement with a larger font claim, “PROVEN TO SAVE UP TO 47% ON YOUR HEATING BILLS!”; about half the subjects shown the stimulus with the disclaimer continued to believe that half or more buyers would save “about 47%” on their heating bills); Kesten C. Green & J. Scott Armstrong, Evidence on the Effects of Mandatory Disclaimers in Advertising, 31 Journal of Public Policy & Marketing 293 (2012) (reviewing evidence); Gita Venkataramani Johar & Carolyn J. Simmons, The Use of Concurrent Disclosures to Correct Invalid Inferences, 26 Journal of Consumer Research 307 (2000) (showing that even subjects who recall a disclosure do not always use it to adjust their judgments).

292 See, e.g., Procter & Gamble Pharmaceuticals, Inc. v. Hoffman-LaRoche, Inc., No. 06-Civ-0034, 2006 U.S. Dist. LEXIS 64363, at *87 (S.D.N.Y. Sept. 6, 2006) (in an unfair competition case, rejecting survey evidence for a host of reasons, including because the disclaimer used was inadequate to construct a control stimulus). Cf. McNeil-PPC v. Pfizer, Inc., 351 F. Supp. 2d 226 (S.D.N.Y. 2005) (rejecting survey evidence where control likely conveyed the same false message conveyed by the tested advertisement); Novartis Consumer Health, Inc. v. Johnson & Johnson-Merck Consumer Pharmaceutical Co., 290 F.3d 578, 599 (3rd Cir. 2002) ("[W]e are skeptical whether disclaimers can cure false advertising claims (made literally or by necessary implication), [however], they may be able to dispel misleading messages implied by a product’s name."); noting that other courts have placed the burden on defendants to demonstrate the efficacy of disclaimers).
influence is being assessed.” However, a problem arises when the “characteristic” whose influence is being assessed is not a single picture or snippet of text but a set of interactions of many aspects of a firm’s website or marketing materials. In many cases, a control is constructed that eliminates some but not all misleading aspects, particularly when deceptive omissions are involved. One survey experiment found that when subjects were shown a telecom’s actual marketing materials (Figure A), 24% understood that there was a charge for cancelling the subscription within two years, 2% gathered that the monthly price would increase after one year, and not a single subject perceived that the monthly fee for premium channels jumps to $47 per month after a 3-month free trial period. With a few minor changes to the advertisement (Figure B; changes marked in fuchsia for illustrative purposes), the percentage of subjects to whom the advertisement correctly conveyed these aspects of the offer doubled (to 47%), increased thirteen-fold (to 26%), and went from 0% to 37%, respectively.

295 Figures copied from Attachment B to the Federal Trade Commission’s Proposed Finding of Fact and Conclusions of Law, Federal Trade Commission v. DirecTV, Inc., Case 4:15-cv-01129-HSG Document 401-3 (filed 09/25/17). Illegible fine print from the bottoms of the figures have been removed, and fuchsia boxes have been added to indicate changes made to create the control stimulus.
CUSTOMER CONFUSION AUDITS: LESSONS FROM THE USE OF CONSUMER CONFUSION EVIDENCE IN THE UNITED STATES

Fig. A: Telecom’s Actual (Test) Advertisement

Fig. B: Telecom’s Modified (Control) Advertisement
Thus, the control stimulus added additional information to the advertisement, but over half of subjects in the control condition continued to, for example, harbor the false belief that there was no cancellation fee for stopping the subscription before the two-year contract period was up. It seems unlikely that U.S. consumers harbor a background belief that all telecom contracts can be cancelled without charge, given the ubiquity of cancellation charges for mobile phone service contracts. Instead, the control appears to have failed to fully remove the confusing elements from the advertisement. The expert herself opined that if she had been free to design an entirely new advertisement, she could have achieved higher levels of consumer understanding.297

For some products or features, it is the design of the product or feature itself that is misleading, such that no control stimulus can eliminate the confusion that the firm creates without changing the design of the product or feature in the control condition. In the same telecom case just discussed, the control condition improved subject understanding by 24 percentage points (from 2% to 26%) of the fact that the monthly price would increase after one year of a two-year contract. But even in the control condition, nearly three quarters of subjects continued to harbor the false belief that the monthly price was fixed for the term of the contract. Perhaps this is because when service contracts have a fixed length, they ordinarily have fixed prices for that time period. Although the telecom’s advertising did not create any such belief, it is possible that through the two-year contract length in the design of the contract, the defendant implied that the price would not change. To fully control for this, a control stimulus might need to change the product itself to eliminate this misleading feature.

A similar issue appears to have been presented in the case of a belt to be worn around the waist that used electrical signals to stimulate involuntary muscle contractions. An advertising onslaught from other similar-looking products made explicit false representations that electrical-stimulation belts would improve abdominal muscle appearance and reduce belly fat. The advertising for the product at issue in the case used the product name (“Abciser”) and photos of well-toned midriffs to obliquely imply that the defendant’s product would have the effects explicitly claimed in the other companies’ advertising. However, without background consumer confusion about the ability of electrical stimulation to improve the midriff, it seems unlikely that the Abciser would have sold well.

The Federal Trade Commission’s expert witness in that case performed a survey experiment in which a control advertisement was created by removing the well-toned midriffs, changing the name of the product, and suggesting that the product could be used for massage. However, a substantial proportion of subjects in the control condition continued to believe that the product would improve abdominal muscles and reduce belly fat.298 In this case, the entire product was likely deceptive, and the very fact that anyone would sell an electrical stimulation waist belt implied the falsehood that such a belt would have a desirable effect. Thus, it is possible that no control could fully eliminate “pre-existing” consumer confusion without changing the entire product being sold.

Potential for Measurement Error. Many of the same sources of measurement error that can undermine the validity of surveys can occur with survey experiments: confusing or leading questions, an interviewer who acts in such a way as to bias results, acquiescence bias, order effects, and issues specific to closed-ended and


298 See In the Matter of Telebrands Corp., 140 F.T.C. 278, 315-327 (2005). The Federal Trade Commission noted, and the expert who conducted the survey experiment agreed, that in creating the control stimulus, it was not possible to remove every misleading element from the defendant’s commercials without fundamentally redesigning the defendant’s commercials. As a result, the difference between the confusion rate found in the test condition and the confusion rate found in the control condition understated the consumer confusion created by the defendant’s commercial.
open-ended questions. These can be mitigated to some extent by proper questionnaire design and fielding, as explained above.

**Potential for Data Analysis Errors.** As is the case for many types of scientific evidence, survey and usability experiment results can suffer methodological defects at the data analysis stage. Again, transparency with respect to the collected data is the best cure.

* * * * *

In sum, properly performed survey and usability experiments have the potential to isolate the causal effect of a firm’s advertising on consumers, at least in the context of twentieth-century marketing. Measurement and data analysis errors can be minimized. As compared to surveys of customer knowledge and beliefs, memory issues pose less of a problem. However, it is not always possible to construct a proper control. Moreover, twenty-first century online marketing and sales practices present looming ecological and population validity challenges to the use of survey and usability experiments for measuring consumer confusion.

Note, however, that when the aim is not to prove that a firm’s promotional materials, website, or disclosures are likely to create customer confusion, survey and usability experiments will continue to be useful. Testing a set of potential “controls,” for example, could allow firms and regulators to isolate problematic features of advertisements, websites, pricing structures and product designs. The intelligence gained from this research could lead to improved firm communications or regulatory action to address the features that are shown to be particularly difficult for consumers to understand.

3. **Copy Tests and Usability Tests (without a Control Condition)**

**KEY CONCEPTS:** Copy and usability tests have many of the same current and incipient validity issues as survey and usability experiments. Without a control condition, they cannot isolate causation, but they also avoid the problem of overcorrecting for legally relevant consumer confusion. They are likely to be particularly valuable when performed by regulators and by firms themselves to identify marketing strategies, website designs, pricing structures, and product and service designs that have the potential to confuse consumers. This data can guide firms and regulators in their efforts to improve consumer understanding of their transactions.

Copy tests and usability tests have been used to gauge consumers’ perceptions of, and confusion about, stimuli relevant to a legal claim. Copy and usability tests are distinct from consumer and customer surveys in that they affirmatively ask participants to derive information from a stimulus during the test, rather than asking respondents to report their existing knowledge or beliefs. Unlike a survey experiment or usability experiment, no control group is tested. However, as can be done in customer surveys, participants can be asked control questions designed to identify participants who are guessing, yea-saying, confused about what the question is asking, or not paying sufficient attention.

Copy tests are a tool originally developed by marketers to help design marketing materials. To perform these tests for litigation purposes, participants are provided with marketing or other communications from the firm and asked questions designed to elicit what messages the participants understand the materials to convey. Sometimes an expert will employ a survey experiment design but will treat responses to open-ended questions as a copy test. That is, for open-ended questions, the respondent confusion rate measured in the control condition is not netted out from the confusion rate measured in the test condition. The Federal Trade Commission has approved of this practice for non-leading open-ended questions.299

---

299 See, e.g., In the Matter of Telebrands Corp., 140 F.T.C. 278, 318 (2005) (“[I]t is appropriate to consider the open-ended responses without netting out any controls.”).
Usability tests are a tool originally used by businesses and others to ensure that their customers or audience can successfully use their websites, electronic applications (“apps”), or other user interfaces. To perform usability testing, participants are typically brought into a lab and asked to perform various tasks using the electronic interface. Participant reactions and actions are recorded to determine whether the interface is unfair or deceptive, such as a website that tricks users into consequences they did not intend or understand. For example, one usability test well-known in the literature brought participants into a lab and asked them to use various electronic tools that claim to limit online behavioral (i.e., personally-targeted) advertising. The research found that consumers had tremendous difficulty using these tools and often erroneously concluded that they had properly configured the tools to block advertising when they had not managed to do so.

• **Construct Validity**

Similarly to survey and usability experiments, copy tests and usability testing assess reactions to the stimuli that are placed before respondents. Thus, copy and usability tests have better construct validity when assessing whether particular text is clearly and conspicuously disclosed or whether a particular communication or website is misleading than when assessing whether a firm is unfairly taking advantage of consumer confusion to make sales or whether a firm has obtained informed consumer consent.

• **Strengths**

**Feasibility.** As with survey and usability experiments, subjects in copy and usability tests are usually taken from the population of consumers who are potential users of the product or service in question. Therefore, subjects are less challenging to locate than if the relevant population were limited to actual customers of a particular product and a particular firm. Further, subjects have generally viewed or used the stimulus just before answering questions, or the stimulus is still present while the respondents are answering questions. Therefore, memory lapses are less of a problem than for customer surveys that take place sometime after the consumer engaged in the relevant decision-making.

• **Weaknesses**

**Looming Impracticability.** For the same reasons discussed above in conjunction with survey and usability experiments and unaided judicial judgment, the proliferation of micro-targeted marketing materials and electronic interfaces may render copy tests and usability experiments impracticable. Testing of representative samples of these stimuli would mitigate this burden, but the sample size needed to produce meaningful results could still be onerous.

**Weak Ecological Validity and Looming Ecological Validity Challenges.** As with survey and usability experiments, copy and usability tests present stimuli in an impoverished context, with fewer and different distractions than those encountered by consumers in the marketplace. Take, for example, a copy test of marketing materials used to sell credit and a credit insurance product. In the marketplace, consumers obtaining a loan might not even realize they had purchased “insurance” with that loan. Copy testing that establishes that consumers who examine marketing materials know they are buying insurance is irrelevant if the problem is that consumers in the real world do not examine the marketing materials.

---

300 For an in-depth understanding of usability testing, see Jonathan Lazar et al., Research Methods in Human-Computer Interaction (Elsevier Science & Technology: 2017).

301 Some usability testing can also be performed in more naturalistic settings. See, e.g., A.J. Brush et al., *In-Home Deployments, in* Studying and Designing Technology for Domestic Life (Tejinder Judge & Carman Neustaedeter, eds., Elsevier: 2014).

For usability tests, it is possible to replicate a website or other electronic interface exactly, but the interface is only part of the environment experienced by users in real life. One usability researcher found that in three lab tests of a particular web-based security system, the average participant completed correctly over 90% of the tasks assigned, but when the system was released in the field, the actual user success rate was closer to 50%. 303

Twenty-first century real-time micro-targeting of marketing and websites presents a further challenge. In the real world, a consumer might be shown one advertisement for a credit card when the consumer is walking into a store and a different advertisement for the same credit card when riding the subway. Similarly, a consumer might be channeled to one website after entering one set of search terms and another website selling the same credit card after entering a different set of search terms. Matching participant environments to the particular advertisements and website designs they would have been shown in those environments is impracticable for an independent researcher to perform.

Population Validity Concerns and Looming Challenges. To be admissible, copy and usability tests must be conducted according to accepted current social science practices. For copy tests, courts have accepted the use of convenience samples, such as consumers at shopping malls. However, it is possible for copy tests to employ techniques that improve population validity, similar to the protocols employed in consumer and customer surveys, survey experiments, and usability experiments. 304

Usability tests for business purposes are often performed on a fairly small number of consumers, and it is generally accepted in the usability literature that fairly small samples produce very useful information for website designers. 305 However, although such small-scale tests can help developers design more usable websites, in a litigation context there is a concern about population validity. Small-scale tests generally lack sufficient statistical power to be used to infer the likely prevalence in the field of usability problems such as misled and confused customers. Usability tests that would produce statistically significant results can be performed. 306 However, to produce statistically significant results, the number of participants needed for in-lab testing could be costly. 307

Another population validity problem for both copy tests and usability tests stems from the micro-targeted nature of twenty-first century marketing and online interfaces. As with survey and usability experiments, although respondents can be selected from a sample frame that is representative of the population of concern overall, matching the correct stimulus or website to each study participant is likely impracticable for independent experts to do.

Weak External Validity due to Subject Disinterest. As explained above in the context of survey and usability experiments, participants might pay less attention to the stimulus in the context of a study than they would in real life, where, for example, their own money is at risk. On the other hand, as also discussed above,

304 For example, a copy test could be fielded using “Google Surveys,” discussed above.
306 As noted above, massive in-field usability experiments are ubiquitous, although these have not been developed to expressly test for confused or tricked consumers, a task for which in-lab usability testing might be better suited.
307 Usability testing recruitment and compensation costs alone were roughly estimated at $125 and $79, respectively, in California in 2003. Costs were a bit lower in other locations but have increased in all locations since then. See Deborah Hinderer Sova & Jakob Nielsen, How to Recruit Participants for Usability Studies 149 & 156 (Nielsen Norman Group: 2003) and https://www.nngroup.com/reports/how-to-recruit-participants-usability-studies/.
participants can sometimes be more attentive during a study than they would be when making decisions in the chaos of daily life.

**Potential for Measurement Error.** As noted, measurement error caused by memory problems is not usually a concern for copy tests and usability tests. However, other sources of measurement error that can undermine the validity of surveys and survey experiments (e.g., ambiguous or leading questions, order effects, yea-saying, or an interviewer who introduces bias or error) can weaken the results of copy tests and usability tests. For copy tests, these can be mitigated by proper questionnaire design and fielding, as also explained above.

For usability tests, there are some different issues. First, to minimize the potentially leading effects of researcher questions on participants, tests can be designed to first observe participants performing tasks and then ask the participants questions about their understanding of the interface. Otherwise, asking a participant about a particular element of the interface could cause the participant to focus on that element more than they would do in normal usage of the interface.308

Second, in some cases usability tests present the risk of an additional type of order effect. Participants can learn from performing one task such that their ability to perform another task on the interface is improved. To identify this and mitigate its effects, usability researchers can engage in counterbalancing, meaning varying the order in which participants perform tasks.309

**Potential for Data Analysis Errors.** Again, data analysis errors can undermine the results of copy and usability tests just as these errors can for consumer or customer surveys, survey experiments, and usability experiments. Transparency can be used to minimize the likelihood of such defects.

* * * * *

It is notable that the trend is for businesses to no longer rely exclusively on consumer copy tests and to instead test marketing, websites, documentation, etc. in the field through actual marketing and sales trials.310 This suggests that copy and usability test evidence ought to be supplemented by field testing, such as through surveys of actual customers.

4. **Expert Facial Analyses and Usability Inspections**

**KEY CONCEPT:** Expert facial analyses and usability inspections are likely to be particularly valuable when performed by regulators and by firms themselves to identify marketing strategies, website designs, pricing structures, and product and service designs that have the potential to confuse consumers and warrant further research.

Expert facial analyses and usability inspections311 consist of expert witnesses applying their experience and knowledge of the relevant principles found in the marketing, psychology, or human-computer interaction (HCI) 

---


310 See, e.g., In the Matter of Telebrands Corp., 140 F.T.C. 278, 282-83 (2005) (noting that the firm’s “‘test’ ads were not simply shown to consumers who participated in focus groups or other types of consumer perception research, but were aired in selected markets for limited periods of time and generated actual sales”); Jeri Smith, *The Perils of Copy Testing in Today’s Advertising Environment*, Quirk Media, July 2016, https://www.quirks.com/articles/the-perils-of-copy-testing-in-today-s-advertising-environment (explaining mismatch between copy testing, the environment in which consumers are exposed to advertising, and the way in which consumers filter and process information today).

311 Usability inspections are also called “heuristic evaluations.” See https://www.usability.gov/how-to-and-tools/methods/heuristic-evaluation.html.
literatures to analyze advertisements, disclosures, websites, and the like.\footnote{Note that the Federal Trade Commission is considered to have its own expertise on consumer perceptions, on which it can rely without reference to outside expert witness testimony. Cf. ECM BioFilms, Inc. v. Federal Trade Commission, 851 F.3d 599, 609 (6th Cir. 2017) ("Because the [Federal Trade Commission] deals continually with deception cases, the Commission is often in a better position than are courts to determine when a practice is ‘deceptive.’"). Nonetheless, social science evidence of consumer perceptions is pervasively used by the Commission in its adjudications.}

“Facial analysis” is the term that is used in the U.S. for analyses of non-interactive print, radio, television, and online communications and “usability inspection” is the term that is increasingly used for interactive interfaces. For both, the expert applies communication and usability principles to the face of the communication or interface at hand, generally without reliance on extrinsic evidence such as copy tests or actual customer behavior.

For example, the communications principles employed by experts to determine when a disclosure is likely to be clear and conspicuous to consumers include the following:

a. Prominence (whether the disclosure is prominent in relation to surrounding materials and visually clear enough for borrowers to notice it and read it);
b. Placement (whether the disclosure is presented in an area where consumers are likely to look);
c. Proximity (whether the disclosure is located in close proximity to the claim being qualified);
d. Avoidability (whether specific techniques have been used to increase the likelihood that consumers will look at the disclosure);
e. Clarity (whether the disclosure is presented in easy to understand language); and
f. Repetition (whether the disclosure is repeated to increase the likelihood of exposure).\footnote{Expressed validity for expert facial analyses and usability inspections is strong for legal claims regarding whether the firm’s materials themselves are likely to create confusion. For claims of unfair practices or lack of informed consent, however, construct validity will depend on whether it is reasonable to conclude that most consumers would notice or understand the relevant material.

Construct validity is an important concern for expert facial analyses and usability inspections because they are used to determine whether a disclosure is likely to be clear and conspicuous to consumers. It is crucial to establish that the principles used by experts are valid and reliable for the purpose of determining whether a disclosure is likely to be clear and conspicuous to consumers.}

These principles are the same for both online and offline communications, although the way in which the principles apply can vary with the format.\footnote{For both, the expert applies communication and usability principles to the face of the communication or interface at hand, generally without reliance on extrinsic evidence such as copy tests or actual customer behavior.}

In most instances, usability inspections will be more complex than other types of expert facial analyses, because the HCI expert must assess an interactive environment. These experts opine on whether and how consumers are likely to perceive not only what immediately appears on a website screen but also content that moves on the screen and content that can only be viewed if the user scrolls, clicks, or hovers. The HCI expert must understand how users are likely to interact with various features of an electronic interface—where users tend to look, what they tend to click, etc.\footnote{In most instances, usability inspections will be more complex than other types of expert facial analyses, because the HCI expert must assess an interactive environment. These experts opine on whether and how consumers are likely to perceive not only what immediately appears on a website screen but also content that moves on the screen and content that can only be viewed if the user scrolls, clicks, or hovers. The HCI expert must understand how users are likely to interact with various features of an electronic interface—where users tend to look, what they tend to click, etc.}—whereas consumers have fewer avenues of interaction with offline content. The HCI expert must also know or determine standard design features for the type of interface under consideration, so as to focus on any features that a user habituated to the standard format might not notice were different on the defendant’s interface.\footnote{In most instances, usability inspections will be more complex than other types of expert facial analyses, because the HCI expert must assess an interactive environment. These experts opine on whether and how consumers are likely to perceive not only what immediately appears on a website screen but also content that moves on the screen and content that can only be viewed if the user scrolls, clicks, or hovers. The HCI expert must understand how users are likely to interact with various features of an electronic interface—where users tend to look, what they tend to click, etc.}

- **Construct Validity**

Construct validity is an important concern for expert facial analyses and usability inspections because they are used to determine whether a disclosure is likely to be clear and conspicuous to consumers. It is crucial to establish that the principles used by experts are valid and reliable for the purpose of determining whether a disclosure is likely to be clear and conspicuous to consumers. For claims of unfair practices or lack of informed consent, however, construct validity will depend on whether it is reasonable to conclude that most consumers would notice or understand the relevant material.

- **Construct Validity**

Construct validity is an important concern for expert facial analyses and usability inspections because they are used to determine whether a disclosure is likely to be clear and conspicuous to consumers. It is crucial to establish that the principles used by experts are valid and reliable for the purpose of determining whether a disclosure is likely to be clear and conspicuous to consumers. For claims of unfair practices or lack of informed consent, however, construct validity will depend on whether it is reasonable to conclude that most consumers would notice or understand the relevant material.

---

\footnote{Note that the Federal Trade Commission is considered to have its own expertise on consumer perceptions, on which it can rely without reference to outside expert witness testimony. Cf. ECM BioFilms, Inc. v. Federal Trade Commission, 851 F.3d 599, 609 (6th Cir. 2017) ("Because the [Federal Trade Commission] deals continually with deception cases, the Commission is often in a better position than are courts to determine when a practice is ‘deceptive.’"). Nonetheless, social science evidence of consumer perceptions is pervasively used by the Commission in its adjudications.}

\footnote{Expressed validity for expert facial analyses and usability inspections is strong for legal claims regarding whether the firm’s materials themselves are likely to create confusion. For claims of unfair practices or lack of informed consent, however, construct validity will depend on whether it is reasonable to conclude that most consumers would notice or understand the relevant material.

Construct validity is an important concern for expert facial analyses and usability inspections because they are used to determine whether a disclosure is likely to be clear and conspicuous to consumers. It is crucial to establish that the principles used by experts are valid and reliable for the purpose of determining whether a disclosure is likely to be clear and conspicuous to consumers. For claims of unfair practices or lack of informed consent, however, construct validity will depend on whether it is reasonable to conclude that most consumers would notice or understand the relevant material.

- **Construct Validity**

Construct validity is an important concern for expert facial analyses and usability inspections because they are used to determine whether a disclosure is likely to be clear and conspicuous to consumers. It is crucial to establish that the principles used by experts are valid and reliable for the purpose of determining whether a disclosure is likely to be clear and conspicuous to consumers. For claims of unfair practices or lack of informed consent, however, construct validity will depend on whether it is reasonable to conclude that most consumers would notice or understand the relevant material.

---

\footnote{In most instances, usability inspections will be more complex than other types of expert facial analyses, because the HCI expert must assess an interactive environment. These experts opine on whether and how consumers are likely to perceive not only what immediately appears on a website screen but also content that moves on the screen and content that can only be viewed if the user scrolls, clicks, or hovers. The HCI expert must understand how users are likely to interact with various features of an electronic interface—where users tend to look, what they tend to click, etc.}
consumers would have acted based on the firm’s materials or would have obtained the relevant information from the firm’s materials. For example, if consumers buy a product based on pre-existing confusion that the firm’s materials do not dispel, the fact that the firm’s materials are not themselves misleading does not mean the firm did not engage in an unfair practice when selling based on the pre-existing confusion. Contrariwise, if consumers are already informed about the key facts of a transaction from another source, the fact that the firm’s materials failed to provide the necessary information does not make the consumers’ consent uninformed.

- **Strengths and Weaknesses**

  Expert facial analyses and usability inspections can be relatively inexpensive and quick to produce, at least when only a few materials or website configurations need to be tested. However, the profusion of varying marketing materials and websites used by firms in the twenty-first century could erase this advantage. An expert inspection of dozens of evolving website designs is likely to be expensive, time-consuming, and quickly outdated.

  Further, as with unaided judicial analysis and copy and usability testing, the expert is examining the particular materials before him or her, without the full context in which the consumer encounters the materials or full background with which the consumer understands the materials, limiting ecological validity.

  Expert facial analyses are likely to more accurately identify how consumers understand business communications than a judicial assessment could do, because the experts are practiced in applying communications and usability principles that a judge typically would not know. Yet, similar to unaided judicial judgment, expert facial analyses and usability inspections can suffer from weak population validity. Experts will not always be familiar with domain or subculture-specific knowledge that consumers of particular products might have that likely affects those consumers’ responses to the particular materials the expert analyzes. To take an example from a case quoted from earlier in this Report, neither a judge nor a communications or usability expert will ordinarily know the beliefs with which a teenage girl approaches a purchase of undergarments; thus both unaided judicial judgment and expert facial analyses of likelihood of confusion for trademark infringement claim purposes in such a case could yield incorrect results.

  Moreover, many communication and usability principles are known at too high a level of generality to predict precisely how they will apply in every situation. In extreme fact patterns, all experts might agree. In subtle cases, different experts could validly come to different conclusions about the likelihood of consumer confusion, making expert facial analyses and usability inspections less useful.

  Despite these weaknesses with respect to using expert facial analyses and usability inspections to prove that a firm’s consumer-facing materials, pricing structures, or product or service designs are deceptive, these analyses and inspections could be valuable for other purposes. Large financial services firms today routinely employ experts to perform this kind of research for a reason—it can quickly flag issues that need additional examination. Just as consumer complaints can serve as an early warning system for regulators, so too expert and usability analyses can identify firms, marketing strategies, website designs, pricing structures, and product and service designs that warrant further research.

  • • • • •

  In sum, the quickly-changing micro-targeted machine-learning-generated marketing of the twenty-first century threatens to undermine the population and ecological validity of survey and usability experiments, copy and usability tests, expert facial analyses, and usability inspections.

  Surveys of actual customers, performed well, have the potential to capture whether actual customers of a particular firm understood themselves to have engaged in a transaction and understood key facts about that transaction, no matter which marketing material each consumer received. Although a survey of customer
knowledge or beliefs cannot isolate whether the firm’s marketing or something else caused customers to be confused, this should be of no consequence in most consumer protection cases—the law prohibits firms from profiting from consumer confusion, regardless of the source of that confusion. Even when the legally-relevant construct is limited to confusion caused by acts or omissions of the firm, in many situations it is reasonable to infer that if a firm’s customers are confused, it is because the firm created that confusion. Customer surveys do present logistical issues for the regulator. If the survey cannot be performed soon after the relevant transaction, customer respondents may need to look at the materials or website they encountered during their interactions with the firm or service provider to aid their recall.

5. A Note on Methodological Imperfections and Convergent Validity

**KEY CONCEPT:** All evidence is fallible. Research employing multiple methodologies can corroborate findings regarding customer knowledge and confusion.

U.S. courts have repeatedly recognized that the scientific method does not produce absolute “truth”—it produces our best approximation of truth. The courts recognize that expert testimony regarding consumer perception is often “more probative than prejudicial” and thus useful to the court even when that testimony is based upon an imperfect methodology imperfectly applied. Some consumer perception evidence is so flawed as to have no probative weight, but evidentiary reliability and thus admissibility does not require perfection. As U.S. Judge Richard Posner opined in one case: “[The] survey [in this case] was not perfect, and this is not news. Trials would be very short if only perfect evidence were admissible.” Demanding “perfection” in social science evidence needlessly increases litigation costs; at some point, any further improvements are not worth the costs.

Even without social science training, many judges can understand, in the context of a particular expert analysis in a particular case, the direction and, very roughly, magnitude in which methodological flaws in evidence collection and analysis are likely to have biased research results. It is thus not uncommon for judges to find that deficiencies in an expert’s work that have been pointed out by the opposing party could have affected the results but would likely have favored that opposing party. Judges also look for logical explanations for divergences between the results of different studies, to select among competing expert opinions.

The most common way for courts to handle deficiencies in social science evidence is to look to corroborating evidence, what social scientists might call multiple measures. If a usability inspection, victim witness testimony, and unaided judicial judgment point in the same direction, each measure thereby has strong convergent validity. Corroborating evidence gives courts confidence that the flaws in any one type of research do not strongly undermine its implications.

---

317 See, e.g., Scott Fetzer Co. v. House of Vacuums, Inc., 381 F.3d 477, 488 (5th Cir. 2004) (“A survey can be so badly flawed that it cannot be used to demonstrate the existence of a question of fact on the likelihood of consumer confusion.”).
318 Indianapolis Colts, Inc. v. Metropolitan Baltimore Football Club, L.P., 34 F.3d 410, 416 (7th Cir. 1994).
320 See, e.g., State v. Vertrue, Inc., 2010 WL 1652448 (noting that customer survey evidence in the case “corroborated evidence presented from several other sources, including inter alia, [the defendant’s] own marketing studies and other internal documents”). Convergent validity is the idea that “different measures of the same hypothetical construct ought to correlate highly with one another if the measures are valid.” Andrew M. Colman, Oxford Dictionary of Psychology (Oxford University Press: 3rd ed. 2014).
V. Conclusion

Consumer knowledge of key facts about the transactions in which they engage is one aim of consumer protection and fair competition law. Knowledge of these facts could help consumers make marketplace decisions that further their own welfare and drive market competition over price and quality. Key facts include whether the consumer has engaged in a transaction, the costs of the transaction, and limitations on the benefits the transaction provides.

Confusion audits and caps are a form of performance-based regulation, regulation that sets outcomes and gives firms the incentive and the flexibility to accomplish those outcomes. Ideally, confusion audits should test of a representative sample of a firm’s actual customers to ascertain whether the firm’s customers understand specified key facts about the transaction they have engaged in with the firm. On the whole, other methodologies for assessing likely customer confusion have more difficult validity problems, but could be more logistically feasible. Confusion caps are limitations on the proportion of a firm’s customers who demonstrate confusion about key facts relating to the transaction they have engaged in with the firm. Reductions in customer confusion can also be pursued through a graduated staircase of caps, to encourage firms to continually reduce their customers’ confusion.

Lessons from the U.S. Experience

U.S. consumer protection and fair competition laws aim to address multiple types of consumer confusion, and therefore the consumer confusion construct that is relevant to a case will depend on the particular legal claim involved. This Report has explained some key features of these laws and the confusion constructs with which they are concerned:

- Some U.S. consumer protection and fair competition laws prohibit creating consumer confusion (deception). Others prohibit profiting from consumer confusion, regardless of the source of that confusion (mis-selling).
- Consumer confusion that is relevant to U.S. consumer protection and fair competition law can include both affirmatively inaccurate beliefs and a lack of knowledge about key facts relating to a transaction.
- When assessing consumer confusion, the consumer knowledge that is legally relevant is knowledge of key transaction facts. To meet the purposes of consumer protection and fair competition law, this consumer knowledge ideally should be of sufficient persistence, precision, and depth to facilitate autonomous consumer decisions such that collectively, consumer choices can induce increased substantive market competition. This optimal persistence, precision, and depth is highly context-specific. However, decreasing consumer confusion is likely to be valuable even if consumer knowledge does not reach optimal levels.
- When enforcement authorities bring actions under consumer protection and fair competition laws, U.S. courts tend to find liability if a “significant minority” of consumers are confused, which courts interpret as meaning at least roughly 10% of consumers.
- The simplest way to set customer confusion rate caps is to start at a high level of confusion and reduce it over time, or to set a graduated staircase of caps and corresponding sanctions.

A variety of epistemologies have been used in U.S. consumer protection and fair competition cases to identify and, in some cases, to quantify, consumer or customer actual or likely confusion. These range from reliance on unaided judicial judgment to the use of modern social-scientific research. Each has strengths and weaknesses:
➢ Unaided judicial judgment is economical but has serious validity problems and could become impracticable as the number of distinct communications and online interface designs produced by a single firm for selling a single product or service increases exponentially.

➢ Customer witness testimony has weak population validity for identifying the presence or absence of confusion about key product or service facts but can provide useful qualitative data.

➢ Firm intent with respect to customer confusion is typically not crucial to claims brought under U.S. consumer protection and fair competition law. Nonetheless, firm behavior and intent evidence can be revealing. However, this type of evidence could become increasingly difficult to obtain as computer algorithms direct more firm marketing and sales activity with less human involvement.

➢ Customer behavior evidence can have strong ecological validity for identifying customer confusion but can have weak population validity.

➢ Customer satisfaction evidence cannot demonstrate an absence of customer confusion about key product or service facts.

In U.S. legal proceedings, social-scientific forms of evidence are usually introduced through expert witnesses. The following are useful to have in mind when considering research performed by expert witnesses in the U.S.:

➢ Expert evidence regarding whether consumers are confused about the key facts regarding their transactions with firms is admissible in U.S. legal proceedings when the research is relevant, is the product of reliable principles and methods, and is performed with the rigor expected of experts in the field.

➢ To produce admissible social-scientific evidence, an expert must have sufficient knowledge and skills to perform the research properly and must have experience performing the same type of research or analysis in academic, business, or government settings. Experts qualified to perform the types of research admissible to demonstrate the presence or absence of consumer confusion hail from a variety of formal fields of training.

➢ Transparency sufficient for the opposing party and the court to fully assess expert evidence is the best way to protect against expert bias or error.

Modern social-scientific research has the potential for greater population validity than most conventional forms of consumer confusion evidence, but each type of research has its own strengths and weaknesses:

➢ Surveys of a firm’s actual customers can have exceptionally strong population validity. These surveys can help assess whether a firm has made unfair or abusive use of existing consumer confusion or has failed to obtain informed consumer consent. Surveys of a firm’s actual customers can also have reasonable construct validity for claims regarding confusion created by the firm. Challenging logistical issues to consider include—how to reach a firm’s actual customers, and how to do so soon enough after the relevant transactions so as to minimize the risk that memory lapses could bias survey results.

➢ Survey and usability experiments can provide evidence that a firm’s specific practices are likely to cause consumer confusion. However, twenty-first century real-time micro-targeting of marketing and online interfaces is likely to undermine the ecological and population validity of these experiments for demonstrating this causation. Further, constructing an adequate control can be challenging. Beyond demonstrating causation, this type of research might productively be used to isolate confusing features of advertisements, websites, pricing structures, and product and service designs, so as to guide firms and regulators in their efforts to improve consumer understanding of their transactions.
Copy and usability tests have many of the same current and incipient validity issues as survey and usability experiments. Without a control condition, they cannot isolate causation, but they also avoid the problem of overcorrecting for legally relevant consumer confusion. They are likely to be particularly valuable when performed by regulators and by firms themselves to identify marketing strategies, website designs, pricing structures, and product and service designs that have the potential to confuse consumers. This data can guide firms and regulators in their efforts to improve consumer understanding of their transactions.

Expert facial analyses and usability inspections are likely to be particularly valuable when performed by regulators and by firms themselves to identify marketing strategies, website designs, pricing structures, and product and service designs that have the potential to confuse consumers and therefore warrant further research.

All evidence is fallible. Research employing multiple methodologies can corroborate findings regarding customer knowledge and confusion.

Looking to the Further Development of Customer Confusion Audits and Caps

Implementing a system of customer confusion audits and caps for regulatory purposes is a context-specific enterprise. It will depend on, e.g.: the nature of the particular customer confusion of concern; the key facts about which customers are confused; the structure of the market with respect to the product or service involved; and the persistence, precision, and depth of understanding that are likely to enhance consumer autonomy and promote substantive competition. Just as firms design their products, marketing, and sales processes and adapt these as contexts change, regulators must test and adapt the specific features of any customer confusion audit and cap system.