

WHITE PAPER

The Super Bowl AI Effect

Why \$8 million in ad spend moves nothing in ChatGPT,
and what it means for brand visibility strategy

AiRR™ Score Research | February 2026

EXECUTIVE SUMMARY

AI-referred consumers convert at 9 times the rate of Google search traffic.¹ In 2025, 70% of daily AI users already made purchases through AI recommendations.² A controlled study of 8 brands across 4 competitive industry pairs found that Super Bowl advertising did not change which brand ChatGPT recommends. In every pair, the brand that led before the game still led one week later.³

The study measured brands at 3 points: before Super Bowl LX on February 8, 2026, the day after the game on February 9, and one week later on February 15. Each brand was scored on where it appears within specific purchase-intent prompts on ChatGPT, producing both a per-prompt brand position and an overall composite score. This is not a measure of search rankings or best answers. It measures precisely where a brand sits when a real customer asks a real question inside ChatGPT, at scale.

Of 8 brands studied, only 1 showed meaningful improvement in AI search positioning over the week following the Super Bowl. Seven brands either held flat or declined. In no case did a Super Bowl ad change the competitive outcome: the brand that ChatGPT recommended before the game was the same brand it recommended after.

The traditional "buzz to rank" flywheel, built for Google, does not transfer to generative AI. Brands cannot buy their way into ChatGPT positioning the same way they purchased SEO prominence.

I. THE SHIFTING SEARCH LANDSCAPE

The Rise of AI-Driven Buying Decisions

Consumer search behavior is undergoing a structural shift. Where buyers once turned to Google to research products and services, a growing segment now turns to conversational AI platforms, primarily ChatGPT but increasingly Claude, Gemini, and Perplexity, to get direct answers and recommendations.

The critical distinction is intent. There are 2 fundamentally different types of search happening today. The first is awareness search: a consumer sees a funny Super Bowl ad, gets curious, and Googles the brand. They are not buying. They are satisfying momentary interest triggered by entertainment. The second is purchase-intent search: a consumer opens ChatGPT and types "what is the best food delivery app right now." They are ready to order. That query is bottom-of-funnel, and the AI's answer determines who gets chosen.

Super Bowl advertising excels at driving the first type of search. The data in this paper shows it does nothing for the second.

The Historical Record

The data on Super Bowl ads and search engagement is unambiguous. EDO measured every ad from Super Bowl LIX and found that T-Mobile's Starlink partnership drove 1,163% more consumer engagement than the average spot, Liquid Death generated 704% more engagement, and RAM drove 8.5x the median ad's response.⁴ The Video Advertising Bureau analyzed 15 brands from Super Bowl LIX and found that

advertisers saw anywhere between a 2x to 50x immediate lift in branded search on game night.⁵

The evidence was confirmed again in 2026. Independent Google Trends analysis of 15 Super Bowl LX advertisers found that some brands hit 6 to 13 times their baseline Google search volume during and after the game.⁶ Hims & Hers, one of the brands tracked in this study, generated a 1,360% spike in Google search interest and was the only brand to produce a second search spike the Monday after the game, driven by controversy around its "Rich People Live Longer" messaging.⁷ The Super Bowl generated an estimated \$550 million in earned media value for brands on social media.⁸

These are Google and social numbers. What happened on ChatGPT was an entirely different story.

Super Bowl ads drive awareness search. ChatGPT purchase-intent queries are where revenue decisions happen. The two channels operate on completely different mechanics.

The Super Bowl is the most-watched broadcast in the world, with Super Bowl LX drawing 124.9 million viewers across NBC, Peacock, Telemundo, and digital platforms.⁹ That scale makes it the strongest possible test case for whether traditional media spend translates to AI positioning. By establishing baselines before the Super Bowl, capturing scores the day after the game, and measuring again one week later, this study tested whether the cultural momentum generated by the game's advertising translated into improved brand positioning within ChatGPT.

II. METHODOLOGY

Study Design

The study tracked 8 brands across 4 competitive industry pairs. Super Bowl LX took place on February 8, 2026. Each pair comprised brands that ran Super Bowl LX ads and compete directly against each other for the same consumer in ChatGPT purchase-intent queries.

Scoring Framework

Each brand was scored across 4 proprietary dimensions that together comprise a composite score: Presence, Perception, Prestige, and Persistence. Scores across all 4 dimensions were aggregated to produce a single number for each brand. Scores range up to 100.

Measurement Points

Brands were measured at 3 points to capture both the immediate and short-term impact of Super Bowl advertising on AI search positioning:

Pre-Game (February 8, 2026): Baseline scores established before the Super Bowl aired. This represents each brand's AI search positioning in the absence of any Super Bowl influence.

Day After (February 9, 2026): Scores captured the day after the game. This tests whether the immediate cultural impact of a Super Bowl ad, the search spikes, social conversation, and press coverage, translated into AI positioning changes.

One Week Later (February 15, 2026): Scores captured the following Sunday, exactly one week after the game. This tests whether any delayed effect emerged as Super Bowl-generated content circulated through the information ecosystem.

Comparing the same day of the week (Sunday to Sunday) controls for any weekly patterns in query behavior or model response consistency.

Analytical Approach

To ensure valid comparisons, each brand was evaluated only against queries relevant to its category. This is a critical methodological point: a shoe brand appearing in a telehealth query would score at the floor by design, not because its AI visibility declined, but because the query is irrelevant to the brand. Raw averages that mix relevant and irrelevant queries produce misleading results.

III. FINDINGS

Finding 1: Super Bowl Advertising Did Not Change Competitive Positioning

The central finding is that in all 4 competitive pairs, the brand that led before the Super Bowl still led one week after. No brand spent its way past a competitor in AI search. The competitive gaps that existed before at least \$8 million in ad spend still existed after it. Because the absolute score movement was negligible for 7 of 8 brands, it is almost certain that their broader category rankings, including against competitors that did not advertise during the Super Bowl, remained unchanged as well.

Finding 2: The Day After Showed Minimal Movement

The morning after the Super Bowl, 6 of 8 brands showed movement of less than 2% in either direction. Two brands showed larger shifts: Hims & Hers rose 16.7%, while Uber Eats declined 3.2%. Neither of these day-after movements changed which brand led its competitive pair.

Finding 3: One Week Later, the Story Was the Same

By February 15, one full week after the game, the competitive picture had not changed. Four brands declined from their pre-game scores: Bud Light (-5.8%), Uber Eats (-3.3%), Michelob Ultra (-3.0%), and Instacart (-1.8%). Three brands showed marginal movement within normal variance: Wix (+0.9%), Squarespace (+1.3%), and Ro (flat). One brand, Hims & Hers, showed meaningful improvement at +27.4%.

Even the Hims & Hers improvement did not change the competitive outcome. Ro led the telehealth pair at 70.89 before the game and 70.87 one week later. Hims & Hers improved from 26.53 to 33.79 but remained more than 37 points behind its direct competitor.

Finding 4: The Google/ChatGPT Divergence

The contrast between channels is stark. Hims & Hers generated a 1,360% spike in Google search interest after its Super Bowl ad, the only brand in this study to produce a second search spike the following

Monday.⁷ On ChatGPT, Hims & Hers was the only brand to show meaningful improvement, gaining 27.4% over the week. But that improvement did not close the gap on its competitor Ro.

The broader pattern held across all advertisers. On Google, most Super Bowl LX advertisers returned to baseline search volume within 8 hours of kickoff, with one analyst noting that an \$8 million ad buys about half a news cycle of attention for most categories.¹⁰ On ChatGPT, the competitive scoreboard did not change.

Finding 5: The Hims & Hers Exception

Hims & Hers was the only brand to show meaningful AI score improvement, and the likely mechanism is instructive. Analytics Mates data showed that Hims & Hers was the only brand out of 15 tracked to produce a second Google search spike the Monday after the game.⁷ The provocative "Rich People Live Longer" messaging generated sustained controversy and press coverage that lasted well beyond the typical 8-hour decay window. This volume of sustained cultural conversation is closer to the kind of signal that influences AI positioning: depth of brand narrative in content that language models can access. A single burst of awareness that fades within hours does not produce the same effect. The week-long controversy the ad generated may have.

IV. THE MECHANISM

How Google and ChatGPT Differ Fundamentally

The traditional SEO flywheel operates on a clear causal chain: advertising drives awareness, awareness drives press coverage and social conversation, that content generates backlinks and signals, and Google's crawler indexes those signals in near-real-time to adjust rankings. The entire system is designed to respond dynamically to new information.

Generative AI models operate on a fundamentally different architecture. A language model's understanding of brands, categories, and recommendations is encoded in its training weights, a static representation of knowledge built from historical data. That knowledge is updated periodically through retraining cycles, not through real-time content indexing. A 30-second ad on February 8th does not update the model's weights. The conversation ChatGPT has with a consumer on February 9th draws on the same brand knowledge it had on February 7th.

Google is a mirror of the present. ChatGPT is a map built from the past. Advertising can change what the mirror reflects. It cannot immediately redraw the map.

The Question of Delayed Impact

One week is a short window. It is possible that Super Bowl advertising generates content and brand signals that take weeks or months to influence AI positioning, particularly if a model retraining cycle is required to incorporate new information into the weights.

This is a reasonable hypothesis, and it is precisely why measurement will continue in the weeks and months ahead. If movement appears at 30, 60, or 90 days, or after a known model update, continuous measurement will detect it. What the data does establish is that the near-term mechanism that makes Super Bowl advertising effective on Google (search spike, content indexing, ranking adjustment) does not operate on ChatGPT within the first week.

The Hims & Hers data offers a possible clue about what does work: sustained cultural conversation that generates a high volume of new content over multiple days, not a single burst of awareness that fades within hours.

What Does Move AI Scores

Ongoing research indicates that AI visibility is driven by factors that accumulate over longer timeframes: sustained category authority, depth of brand narrative in training-eligible content, consistent association with specific use cases and consumer problems, and third-party validation in authoritative sources. These are signals that build over months and years, not days.

This has practical implications for marketing strategy. AI visibility is not a media buy. It is a content and authority strategy, more analogous to brand-building and earned media than to paid amplification. Brands that score well have typically established deep, consistent associations with their category in the kinds of content that language models are trained on.

V. IMPLICATIONS FOR CMOS

AI Has Become a Blind Spot in Marketing Measurement

The marketing measurement stack has expanded significantly over the past decade. CMOs now have granular visibility into paid search ROI, organic search share of voice, social engagement and sentiment, programmatic performance, email conversion, and attribution across the full funnel. Yet none of these instruments measure the channel where a growing percentage of bottom-of-funnel decisions are being made.

The Super Bowl data makes this blind spot concrete. Brands spent tens of millions of dollars on advertising during a single event. Their AI visibility either stayed flat or declined. No measurement system in a standard marketing stack would have detected this. No budget line was allocated to understand it.

The stakes of this blind spot are rising fast. A January 2026 survey of over 1,000 U.S. consumers by PartnerCentric found that 64% plan to use AI chatbots for shopping in 2026, and nearly 1 in 4 plan to make AI their default way to shop.¹¹ Among daily AI users, 70% already made purchases through AI recommendations in 2025, spending an average of \$540 across 9 transactions.¹²

Separately, data from Seer Interactive found that ChatGPT search traffic converts at 15.9% compared to Google's organic rate of 1.76%, meaning AI-referred visitors are 9 times more likely to buy.¹ McKinsey projects \$750 billion in U.S. consumer spending will flow through AI-powered search by 2028.¹³

The consumer is already making purchase decisions inside ChatGPT. The question is whether your brand shows up when they do.

The AI Visibility Gap Is a Competitive Risk

In categories where AI-generated recommendations are increasingly influencing purchase decisions, a brand's score relative to its direct competitor determines which brand gets recommended first. The study's competitive pairs reveal gaps that persisted through the Super Bowl regardless of advertising spend.

In website builders, Wix leads Squarespace by 10.6 points before the game and 10.4 points one week later. In telehealth, Ro leads Hims & Hers by 44.4 points before the game and 37.1 points one week later. In beer, Bud Light leads Michelob Ultra by 7.1 points before the game and 4.9 points one week later. In delivery, Instacart leads Uber Eats by 8.1 points before the game and 8.8 points one week later.³ Every leader held.

To understand what these gaps mean in practice, consider the scale of the channel. ChatGPT processes 2.6 billion messages every day, across 188 countries, with over 800 million weekly active users.¹⁴ A meaningful share of those queries are purchase-intent questions in categories exactly like the ones measured in this study: which delivery service should I use, which telehealth platform is better, which website builder is right for my business.

In each of those moments, ChatGPT typically surfaces 3 to 4 brands in its answer. The brand with the higher score captures that recommendation. The brand with the lower score does not. At scale, across millions of queries every day, the competitive gap between brands in AI search is not a dashboard metric. It is the difference between being the answer and not being in the conversation.

Moving up is not an incremental gain. It is a compounding shift in how often your brand gets recommended at the exact moment a consumer is ready to decide. Moving down, or standing still while a competitor improves, is the equivalent of losing shelf space in a store that never closes.

Measurement Must Precede Strategy

The practical starting point for any AI visibility strategy is measurement. Without a baseline score across all 4 dimensions, there is no basis for strategic prioritization, no ability to track the impact of content investments on AI positioning, and no way to benchmark against competitors.

ANALYSIS

Steven Perlman

Every brand in this study was told, correctly, that their Super Bowl ad would generate millions of Google searches. What they may not have realized is that those searches would have limited, if any, immediate impact on which brand ChatGPT recommends when a consumer is ready to buy.

The competitive finding is the headline. In all 4 industry pairs, the brand that led before kickoff still led the day after kickoff as well as one week later. Wix over Squarespace. Ro over Hims & Hers. Bud Light over Michelob Ultra. Instacart over Uber Eats. At least \$8 million per brand, and the scoreboard did not change.

And because the absolute AiRR Score movement was negligible for 7 of the 8 brands, it is almost certain that their rankings against competitors who chose not to spend \$8 million on a Super Bowl ad remained unchanged as well. ChatGPT has become the Super Bowl of online purchasing decisions, where millions of consumers go every day to decide what to buy, and the brands competing in it need to know where they stand.

The Hims & Hers data is the most instructive result in the study, and it raises a fair question: is it an outlier? Consider the facts. Hims & Hers was the only brand out of 15 Super Bowl advertisers to generate a second Google search spike the Monday after the game. It entered the study with the lowest AiRR Score of any brand tracked, at 26.53, roughly half the next lowest brand (Uber Eats at 57.69). And its 27.4% AiRR Score improvement, while meaningful, still left it more than 37 points behind its competitor Ro. The score may have been artificially low to begin with, meaning the improvement reflects a correction rather than a Super Bowl effect. Or the sustained controversy genuinely moved the needle for 1 out of 8 brands. Either way, 1 out of 8 is not a compelling ROI for an \$8 million investment.

This paper is not an argument against Super Bowl advertising. The Super Bowl is the most-watched broadcast in the world, and its ability to generate awareness, cultural conversation, and Google search volume was confirmed again in 2026. But the value of that Google search boom is diminishing. Gartner projects that traditional search engine volume will drop 25% by 2026 as consumers shift to AI assistants.¹⁶ Google's U.S. market share has already declined from approximately 92% to under 86% since ChatGPT launched.¹⁷ Meanwhile, 64% of Google searches now end without a single click to an external website.¹⁸ The channel that Super Bowl ads reliably activate is shrinking in reach and effectiveness at the same time that AI-driven purchasing is growing. The measurement framework for evaluating a Super Bowl investment must account for both trends.

CMOs now face a measurement challenge that did not exist 2 years ago. Their existing tools measure 6 channels. Now AI search is the seventh. No line item in the marketing budget is allocated to measuring it. No dashboard tracks it. The Super Bowl data proves this is no longer a theoretical concern. Brands are spending at the highest levels of the media market and generating zero immediate movement in the competitive landscape of a channel that converts at 9 times the rate of Google.¹

Knowing where your brand stands in AI search, how far you are from your competitor, and what it takes to move. That is the starting point. That is what AiRR Score measures.

About AiRR Score

The data in this paper was collected using AiRR Score, the first system of record for how brands appear within generative AI platforms. AiRR Score measures brand visibility and positioning within ChatGPT across purchase-intent queries that drive real consumer decisions.

AiRR Score is pioneering a new measurement category. Using a proprietary 4P framework measuring Presence, Perception, Prestige, and Persistence, AiRR Score tells a brand not just its overall position within AI search, but how it ranks against direct competitors across specific high-value prompts at scale.

The Super Bowl proved that AI search visibility requires a different playbook. The first step is knowing the score.

airrscore.com

SOURCES

1. Seer Interactive, "AI Platform Conversion Analysis," June 2025. ChatGPT converts at 15.9% compared to Google's organic rate of 1.76% (9.03x). Based on 11,000 AI-referred sessions analyzed.
2. PartnerCentric, January 2026. Survey of 1,000+ U.S. consumers. 70% of daily AI users made purchases through AI recommendations in 2025, spending an average of \$540 across 9 transactions.
3. AiRR Score Research, February 2026. Controlled study of 8 brands across 4 industry pairs, February 8-15, 2026. Baseline established February 8 before game. All scores and methodology documentation available under NDA.
4. EDO, Super Bowl LIX Ad Rankings, February 2025. T-Mobile: 1,163% more engagement than average spot. Liquid Death: 704%. RAM: 8.5x median ad response. Based on branded search and site visit engagement immediately following each airing.
5. Video Advertising Bureau, "Super Bowl 2025 Advertiser Success Stories," February 2025. 15 brands analyzed. Branded search lift ranged from 2x to 50x on game night.
6. Analytics Mates, February 2026. Google Trends analysis of 15 Super Bowl LX advertisers. Some brands reached 6x to 13x baseline Google search volume.
7. Analytics Mates, February 2026. Hims & Hers generated a 1,360% Google Trends spike and was the only tracked brand to produce a second search spike Monday morning, driven by controversy around its "Rich People Live Longer" messaging.
8. Marketing Dive, February 2026. Super Bowl LX generated estimated \$550 million in earned media value on social media.
9. ALM Corp, February 2026. Super Bowl LX drew 124.9 million viewers across NBC, Peacock, Telemundo, and digital platforms. AdImpact reported streaming accounted for 56% of the audience, a first for any Super Bowl.
10. Analytics Mates, February 2026. Most Super Bowl LX advertisers returned to baseline Google search volume within 8 hours of kickoff.
11. PartnerCentric, January 2026. 64% of consumers plan to use AI chatbots for shopping in 2026. Nearly 1 in 4 plan to make AI their default way to shop.
12. PartnerCentric, January 2026. Among daily AI users: 70% made purchases through AI recommendations, average \$540 across 9 transactions.
13. McKinsey, 2025. Projects \$750 billion in U.S. consumer spending through AI-powered search by 2028.
14. NBER Working Paper 34255 (Chatterji, Cunningham, Deming et al., Harvard/NBER, September 2025). ChatGPT: 2.6 billion daily messages, 188 countries. 800 million weekly active users per Sam Altman, OpenAI Dev Day, October 2025.

15. NBC charged a base rate of \$10 million for a 30-second commercial during Super Bowl LX (Wikipedia / NBCUniversal). Early negotiations started at approximately \$7 million (Marketing Dive, December 2025), with industry sources reporting an \$8 million average price point (ALM Corp, February 2026).

16. Gartner, 2025. Projects traditional search engine volume will drop 25% by 2026 as users shift to generative AI assistants and chatbots.

17. GetTheClicks / StatCounter, October 2025. Google's U.S. search market share declined from approximately 92% to 85.67% following the launch of ChatGPT.

18. SparkToro / Datos, 2025. 64% of Google searches end without a click to an external website. Zero-click searches exceeded 3 trillion in 2024.