

THE ORDER × AUTOPILOT

The Third Shelf

The Alexa for Shopping Atlas — to our knowledge the first large-scale study of how Amazon's AI shopping assistant recommends products versus the organic and sponsored search results.

NON-BRANDED QUERIES

1,963

DISTINCT PRODUCTS

70,176

AI RECOMMENDATIONS

13,177

SEARCH LISTINGS

109,244

EXECUTIVE SUMMARY

Amazon now has a third shelf. It plays by its own rules.

Amazon's AI shopping assistant — **Alexa for Shopping**, built on the Rufus engine Amazon reports assisted over 300 million customers in 2025 — recommends a strikingly different set of products than Amazon's own search results. This study captured the assistant's recommendations side-by-side with the full first search page for 1,963 non-branded shopping queries. Five findings follow.

- 1 The AI decouples from organic rank.** 63.9% of its picks fall outside the organic top-10 (robust at 63–65% across every specification tested); 40.9% are not matched anywhere on the captured page (organic or sponsored).
- 2 Ads overlap winners.** Ads appear on 14.3% of picks, but 83% of those products already rank organically; only 2.4% of picks are advertised without any organic ranking — a pattern consistent with ads accompanying organic strength rather than substituting for it.
- 3 It anchors, then diversifies.** The assistant's first pick matches the organic top-10 55.4% of the time, decaying to ~12% by its deepest slot — later picks go deeper on the page and further off it.
- 4 Categories differ sharply.** Apparel and long-tail queries are 75–80% decoupled from the top-10; Beauty and Health & Supplements remain closest to traditional search (~54%).
- 5 The fundamentals still matter — as correlates, not levers.** Organic rank, sales velocity and review depth are the strongest correlates of being picked; an Amazon's Choice badge adds no positive independent signal once those are controlled.

A note on method: this study corrected itself. An internal adversarial audit led us to retract an early “winner-take-most” brand-concentration claim (a measurement artifact) and to widen stated uncertainty. **The numbers here are the corrected, conservative ones.**

METHODOLOGY

Side-by-side capture, **verified twice**

The design measures **category discovery** — how products surface when a shopper describes a need — not brand lookups.

THE PIPELINE

- 1 Query bank.** 1,963 non-branded queries drawn from real Amazon search-demand data (Search Query Performance), stratified across five category groups; branded terms removed. Each query is tagged by category and intent type (broad, narrow, transactional, problem-framed, comparative).
- 2 Same-session capture.** Each query was run once through the assistant (US, May–June 2026), capturing its full recommendation list alongside the first search page as captured — a median 60 tiles: ~48 organic ranks + ~12 sponsored placements, contiguous as rendered.
- 3 Crosswalk.** 12,810 of the 13,177 recommendations (queries with failed page captures excluded) were matched to their exact position in the same-session results — organic rank 1–48, sponsored, or beyond the captured page.
- 4 Statistics.** Wilson intervals; cluster-robust inference (picks nest within queries; design effect 1.36×); an item-count-matched permutation test for brand concentration; and a logistic model with cluster-robust errors, collinearity checks and out-of-sample validation (AUC 0.81).
- 5 Adversarial self-audit.** Independent recomputation of every headline figure from the raw data (34/34 checks confirmed) plus an interpretation critique that tightened labels and retracted one finding.

EXAMPLE QUERIES FROM THE BANK

Broad — “steam mop”, “womens sandals”, “collagen”

Narrow — “tv stands for living room”, “steel toe boots for men”

Transactional — “home decor clearance”

Problem-framed — “jar openers top rated for seniors”

NAMING & SOURCES

The assistant surface was captured (May–June 2026) under its then-current brand, **Rufus**; Amazon has since unified it with Alexa+ personalization under **Alexa for Shopping** (Amazon announcement: aboutamazon.com/news/retail/alexa-for-shopping-ai-assistant — also the source of the 300-million-customers figure cited in the summary). This paper refers to the assistant as Alexa for Shopping throughout. The study was designed and analyzed by Autopilot; The Order co-publishes.

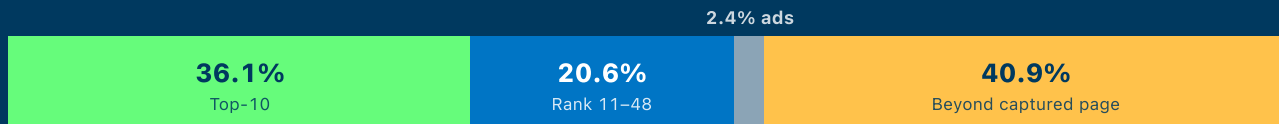
Boundary to keep in mind: the organic capture reaches ~rank 48 (the visible page holds ~60 tiles). **“Beyond the captured page” is an upper bound** — an inseparable mix of the page-1 tail, page 2+, and non-ranking items. It does not mean “not for sale.”

FINDING 1

63.9% of AI picks sit outside the organic top-10

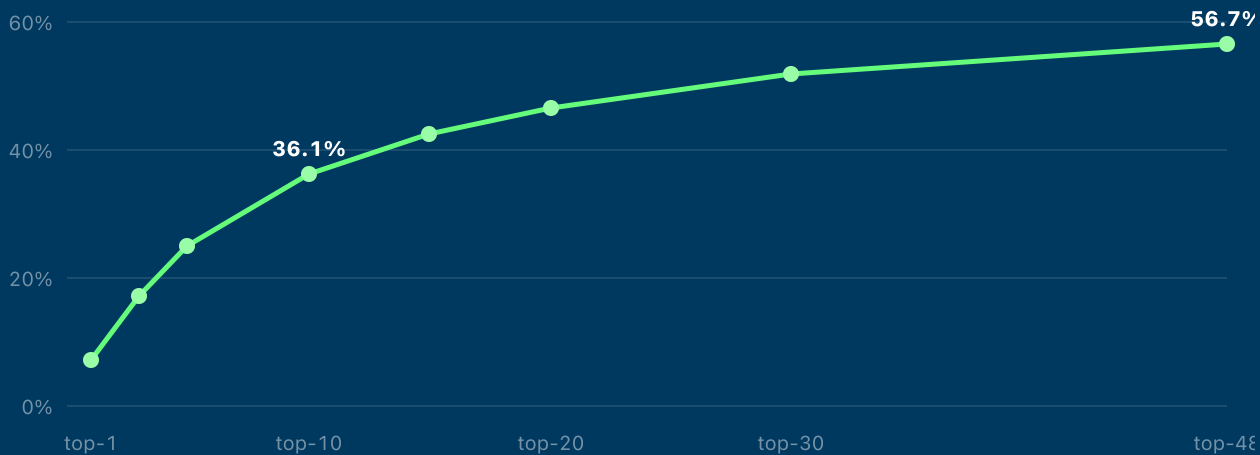
Across 12,810 picks from 1,907 valid-capture queries (95% CI 63.1–64.8%; 63.1% after clustering by query, CI 62.0–64.3%). The result is robust at 63–65% across every specification tested — including keeping the excluded failed captures (64.9%).

WHERE EVERY PICK SITS — THE FOUR-WAY SPLIT



Where 12,810 Alexa picks sit relative to the captured page-1 results, organic and sponsored (valid-capture queries)

WIDENING THE WINDOW HAS FAST-DIMINISHING RETURNS

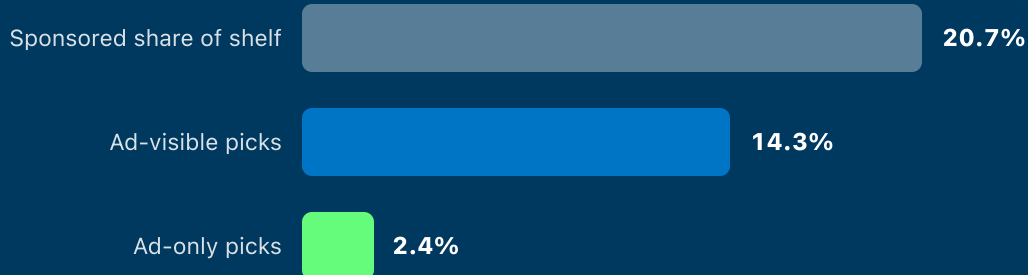


Cumulative share of all picks matched within the top-K organic ranks. Top-10 catches 36.1%; the full captured window only 56.7% — so $\geq 40\%$ of picks require an explanation other than first-page organic position. The decoupling is not an artifact of the top-10 cutoff.

- One-fifth of all picks (**20.6%**) sit deeper on the same captured page — rank 11–48, median rank **20**.
- **40.9%** are beyond the captured organic window entirely.
- Implication: **organic-rank dashboards systematically under-count** a brand's true AI presence.
- Ordering is also weakly linked to rank across the full depth (Spearman $\rho = 0.26$).

FINDING 2

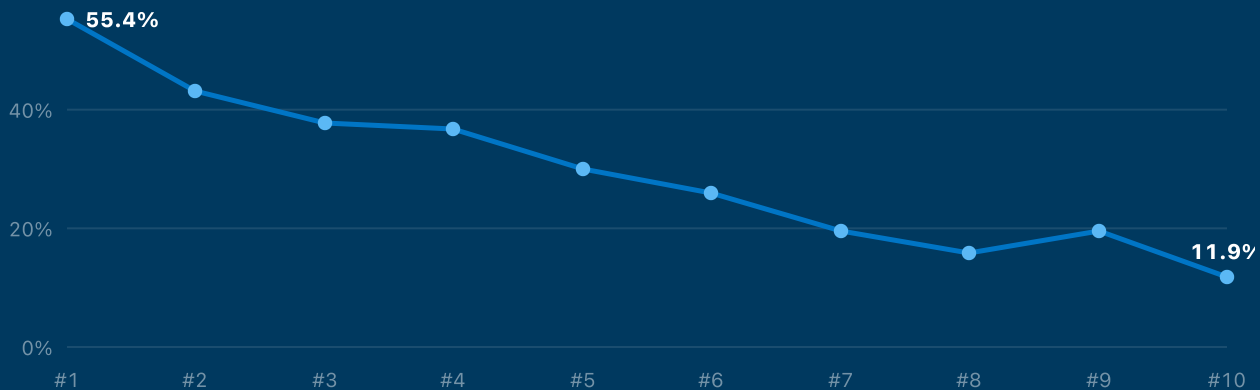
Ads overlap winners — they rarely appear alone



Sponsored placements are 20.7% of page-1 tiles. Ads appear on 14.3% of the assistant's picks — but **83% of those products already rank organically** (mostly top-10). Only 2.4% of picks are advertised with no organic ranking. In rate terms: an ad-only product is picked 1.9% of the time, versus 24.5% for products that both rank and advertise — a difference consistent with organic strength rather than ad effect; this observational design cannot isolate the ad's own contribution.

FINDING 3

It anchors on an incumbent, then diversifies



Share of each AI pick slot matching the organic top-10. The first slot behaves like an incumbent position (55.4%); by slot 8 the median on-page pick sits at organic rank 16, and the share of picks beyond the captured page rises from 28% (slot 1) to 58% (slot 8). The decline is near-monotonic (trend $z = -24.9$); the slot-9 uptick (19.6%, $n=97$ picks vs ~1,900 for slots 1–6) is consistent with sampling noise.

FINDING 4

Most open where rank matters least



Share of picks outside the organic top-10, by category ($\chi^2 = 470$, Cramér's $V = 0.19$). Apparel and long-tail "Other" queries are the open arena; Beauty and Health & Supplements stay closest to traditional search — decoupling is lowest there.

FINDING 5

What correlates with a pick — and what doesn't

- **Organic rank** is the dominant correlate (adjusted OR 0.914 per worse rank; picked median rank 7 vs 25).
- **Review depth**: ~3.5× more reviews (median 6,595 vs 1,894).
- **Brand concentration is small**. An early "winner-take-most" reading collapsed under an item-count-matched test to a +0.055 excess on the concentration index, and was revised accordingly.
- **Sales velocity**: picked products carry a ~3.3× higher "bought last month" figure (Amazon's on-tile label — a rounded proxy, not exact sales).
- **Amazon's Choice adds no positive independent signal** once rank, reviews and velocity are controlled (if anything marginally negative: OR 0.87, CI 0.81–0.93).
- All of these are **associations, not proven causes** — being recommended itself drives sales and reviews.

PART TWO

Five queries, picked apart

Aggregates show the pattern; single queries make it concrete. The next five pages take one real, non-branded query from each category group and map **every product the assistant recommended** against the full first-page shelf captured in the same session — its exact organic rank, and whether it was advertised. Product names and figures are shown as captured from the public search page.

HOW TO READ THE MAPS

- 1 Pick number · green = in organic top-10
 - 2 White = deeper on page (rank 11–48)
 - 3 Amber = beyond the captured results (page-1 tail, page 2+, or not ranking — inseparable)
- + AD** Product also appeared as a sponsored placement
-  Dot position = organic rank on a 1→48 scale; right zone = beyond the captured results

THE AGGREGATE BACKDROP FOR EACH CATEGORY

CATEGORY	PICKS OUTSIDE TOP-10	BEYOND CAPTURED WINDOW	MEDIAN ON-PAGE RANK
Home & Kitchen	60.0%	33.7%	7
Apparel	75.0%	57.6%	7
Beauty	54.3%	27.4%	6
Health & Supplements	53.6%	28.3%	6
Other (long-tail)	80.2%	63.4%	9

These five queries were chosen to be representative, not extreme. **Three of the five anchor on a top-10 product** before diversifying; two lead with picks beyond the captured window — the same mix the aggregate numbers describe.

Want more than five? The free companion volume — **The Third Shelf — Deep-Dive Library** — maps 683 queries in exactly this format and indexes 903 brands to the pages where they appear. See if your brand is on the list: autopilotbrand.com/thirdshelf.

CATEGORY DEEP DIVE · HOME & KITCHEN

"steam mop"

One real, non-branded query from the study. Every product the assistant recommended, mapped against the page-1 shelf captured in the same session (48 organic ranks + 12 sponsored placements).

EVERY PICK, MAPPED TO ITS ORGANIC RANK



WHAT THIS QUERY SHOWS

- Alexa anchors on the organic **#1**, then reaches to ranks **16** and **31**.
- **2 of 7** picks are not matched anywhere in the captured results — a blind spot for rank-based tracking.
- Only **1 of 7** picks was running an ad (and it also ranked #5 organically).

3 of 7 picks in organic top-10

2 deeper on page (rank 11–48)

2 beyond captured results

1 advertised

Across all Home & Kitchen queries in the study: **60.0% of picks fall outside the organic top-10** · 33.7% are beyond the captured window · median on-page rank 7.

CATEGORY DEEP DIVE · APPAREL

"hat rack"

One real, non-branded query from the study. Every product the assistant recommended, mapped against the page-1 shelf captured in the same session (**48 organic ranks + 12 sponsored placements**).

EVERY PICK, MAPPED TO ITS ORGANIC RANK



WHAT THIS QUERY SHOWS

- The AI's **lead pick** sits at organic **#9** — while the organic **#1** appears only at pick **#7**.
- Picks come from ranks 1, 4, 9, 23 and 25 — the AI's order is not the ranking.
- Both advertised picks also ranked organically (#4, #25) — ads accompanied existing listings.
- *Categorized under Apparel per the study's query-bank taxonomy (headwear accessories & organizers).*

3 of 7 picks in organic top-10

2 deeper on page (rank 11–48)

2 beyond captured results

2 advertised

Across all Apparel queries in the study: **75.0% of picks fall outside the organic top-10** · 57.6% are beyond the captured window · median on-page rank 7.

CATEGORY DEEP DIVE · BEAUTY

"eye cream"

One real, non-branded query from the study. Every product the assistant recommended, mapped against the page-1 shelf captured in the same session (48 organic ranks + 12 sponsored placements).

EVERY PICK, MAPPED TO ITS ORGANIC RANK



WHAT THIS QUERY SHOWS

- The AI's **#1 recommendation is beyond the captured results** (as is pick #6).
- The organic **#2** product appears — but in the AI's **last** slot.
- 4 of 7 picks are top-10 products, yet the AI reshuffles their order entirely.

4 of 7 picks in organic top-10

1 deeper on page (rank 11–48)

2 beyond captured results

2 advertised

Across all Beauty queries in the study: **54.3% of picks fall outside the organic top-10** · 27.4% are beyond the captured window · median on-page rank 6.

CATEGORY DEEP DIVE · HEALTH & SUPPLEMENTS

"collagen"

One real, non-branded query from the study. Every product the assistant recommended, mapped against the page-1 shelf captured in the same session (**48 organic ranks + 12 sponsored placements**).

EVERY PICK, MAPPED TO ITS ORGANIC RANK



WHAT THIS QUERY SHOWS

- A textbook blend: picks from ranks **2, 3, 7, 9, 12** — plus rank **24** and one beyond the captured results.
- Five of seven picks carry **46k–214k reviews**; the two below that range sit at 21k and 6k — the pick beyond the captured window itself carries 73k.
- Only **1 of 7** picks was advertised (it also ranked #3) — six appeared without a same-session ad.

4 of 7 picks in organic top-10

2 deeper on page (rank 11–48)

1 beyond captured results

1 advertised

Across all Health & Supplements queries in the study: **53.6% of picks fall outside the organic top-10** · 28.3% are beyond the captured window · median on-page rank 6.

CATEGORY DEEP DIVE · OTHER

"jar openers top rated for seniors"

One real, non-branded query from the study. Every product the assistant recommended, mapped against the page-1 shelf captured in the same session (47 organic ranks + 13 sponsored placements).

EVERY PICK, MAPPED TO ITS ORGANIC RANK



WHAT THIS QUERY SHOWS

- The AI's **top three picks are all beyond the captured results** — it reinterpreted the need and led with electric can openers.
- The jar openers it did pick rank **1, 6, 8** and **30** organically.
- A query written in shopper-intent language ("for seniors") surfaced products a top-10 rank tracker would not show.

3 of 7 picks in organic top-10

1 deeper on page (rank 11–48)

3 beyond captured results

1 advertised

Across all Other queries in the study: **80.2% of picks fall outside the organic top-10** · 63.4% are beyond the captured window · median on-page rank 9.

READ RESPONSIBLY

What this study **does not** establish

The findings are observational and reported with their boundaries. Five limits matter most.

- 1 **Single-run snapshot.** One capture per query, US marketplace, May–June 2026. Day-to-day stability and personalization are not measured.
- 2 **Associational, not causal.** The assistant's ranking function is unobserved and reverse causation is live — being recommended drives sales and reviews. “Correlates” are hypotheses to test, not proven levers.
- 3 **“Beyond the page” is an upper bound.** Organic capture reaches ~rank 48 of a ~60-tile page; unmatched picks may sit in the page-1 tail, on page 2+, or not rank at all.
- 4 **Attributes cover the visible half.** Product-level attributes exist only for the 56.7% of picks matched on the captured page; the decoupled remainder has no attribute data.
- 5 **Proxy metrics.** “Sales velocity” reads Amazon's rounded “X+ bought last month” label (shown on ~83% of tiles), not exact unit sales.

Everything in this paper survived independent recomputation from the raw capture data and an adversarial interpretation review. **Where a finding didn't survive, we said so.**

FOR BRANDS

Four moves that follow from the evidence

- 1 Instrument the shelf.** Track which AI picks are yours, separately from organic rank — rank dashboards under-count AI presence by design.
- 2 Write for shopper intent.** In our sample, queries phrased as needs (“top rated for seniors”) surfaced different products than keyword-style queries — making intent-rich listing language the natural hypothesis to test.
- 3 Re-examine ad spend aimed at AI visibility.** Advertised-but-unranked products accounted for just 2.4% of picks — the data does not support “ads earn AI picks” as a working assumption.
- 4 Pick your battles.** The most decoupled categories are where rank incumbents are most exposed — and where challengers have the clearest opening.

See your brand's third shelf

The per-query maps in Part Two can be produced for any brand: your ~100 core discovery queries, your share of AI recommendations relative to your organic and paid presence, and where competitors are picked instead. For a worked example or a brand-level analysis: christian@autopilotbrand.com · autopilotbrand.com

Companion volume: The Third Shelf — Deep-Dive Library. 683 queries mapped, 903 brands indexed — free at autopilotbrand.com/thirdshelf. The Third Shelf is an ongoing study — subsequent waves will track how this shelf shifts.



ABOUT THE PUBLISHERS

Autopilot (autopilotbrand.com) builds software and services for brands on Amazon: content produced from a brand's broader context, ongoing optimization informed by intent language, search-query and advertising data, reviews and DTC insights, and market-share-based measurement of results.

The Order (the1order.substack.com) co-publishes this research. The study itself is vendor-neutral; the methodology and supporting data are available on request for verification.