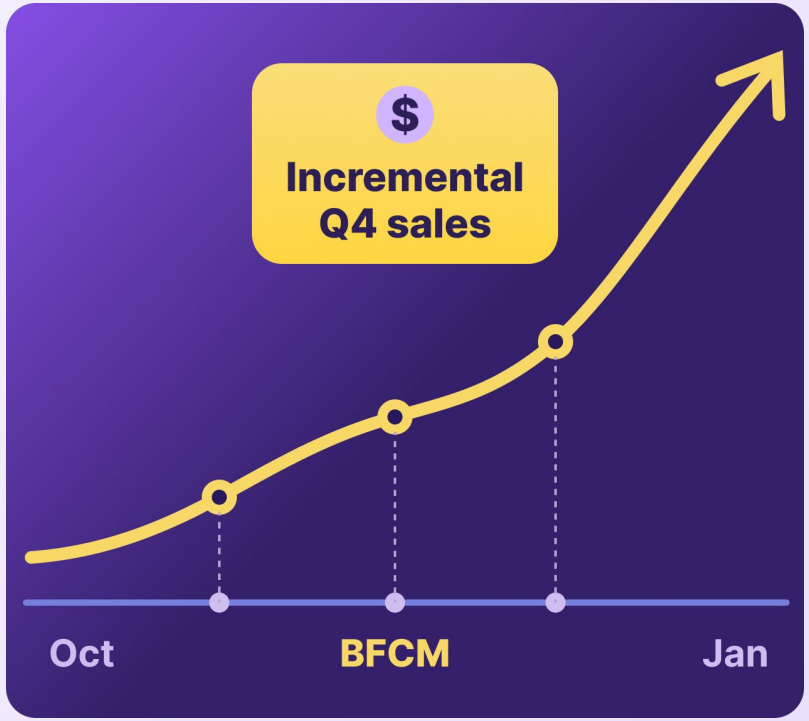
WorkMagic

Playbook: Winning Q4 - Your Most Important Season

Plan for holiday growth, then let lift tests show which dollars earned it

What WorkMagic incrementality tests show about measuring, testing and budgeting through peak season.



Market Insights 2026 · WorkMagic Research

What's Inside

Three sections, each answering a question brands and partners ask before the holidays

01

Why Peak Season Breaks Measurement

Should we measure and test in Q4 at all?

- Holiday growth and the efficiency dip
- Where attribution and lift disagree

02

Scoring and Funding the Season

What does a good holiday look like, and where should the next dollar go?

- New customers, wherever they buy
- Bigger budgets, different returns
- From holiday test to holiday budget

03

Building Your Holiday Test Plan

How should we test in Q4, and what are the risks?

- Lock measurement before peak
- Holdouts that don't cost the season
- Which test for which channel

Plus holiday testing myths, brand spotlights and key takeaways.

Copyright © 2026 WorkMagic

The Data Behind Our Q4 Insights

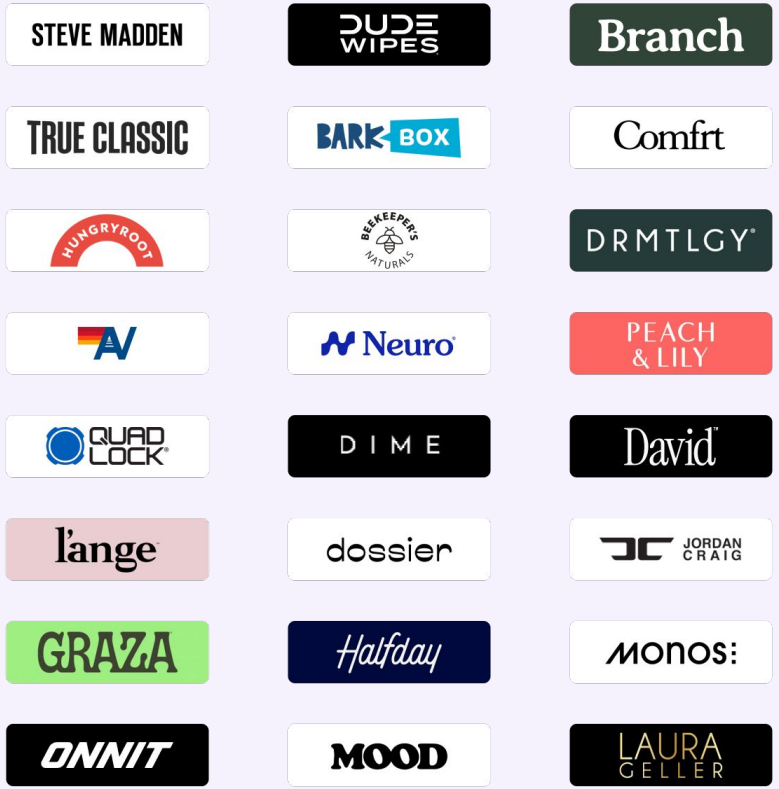
A breakdown of all the tests we've ran and analyzed to power our Q4 Playbook

298 incrementality test groups run in 2025

87 advertisers tested across the year

94 test groups started in Q4 2025

53 advertisers tested during peak season

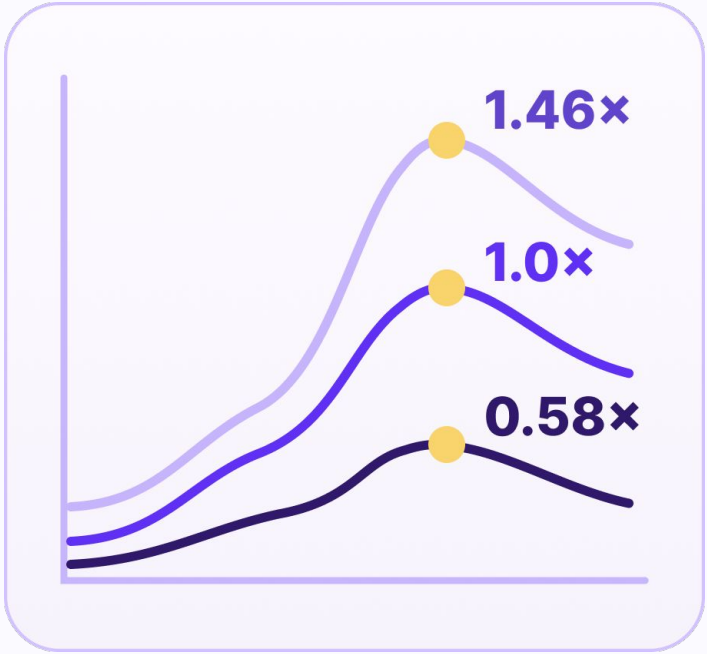


Copyright © 2026 WorkMagic

SECTION 01

Why Peak Season Breaks Measurement

Holiday sales rise with or without your ads, and non-incremental measurement isn’t smart enough to keep up.



AD SPEND

Median monthly spend of advertisers in Q4

+26%

MONTHLY SALES

Median monthly sales of brands in Q4

+28%

MER

Median sales per ad dollar

–5%

Copyright © 2026 WorkMagic

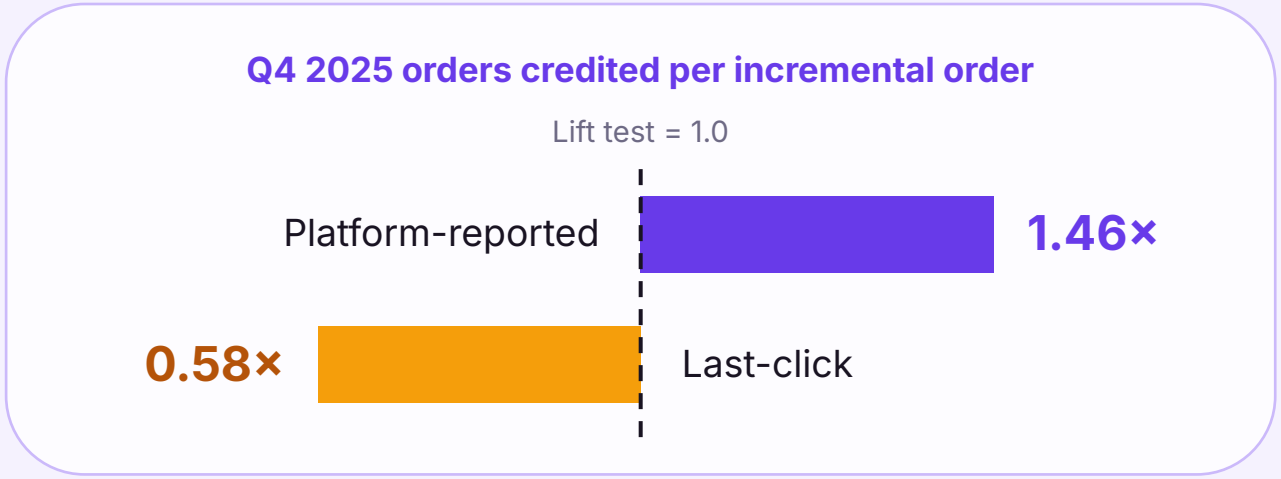
Holiday growth & efficiency

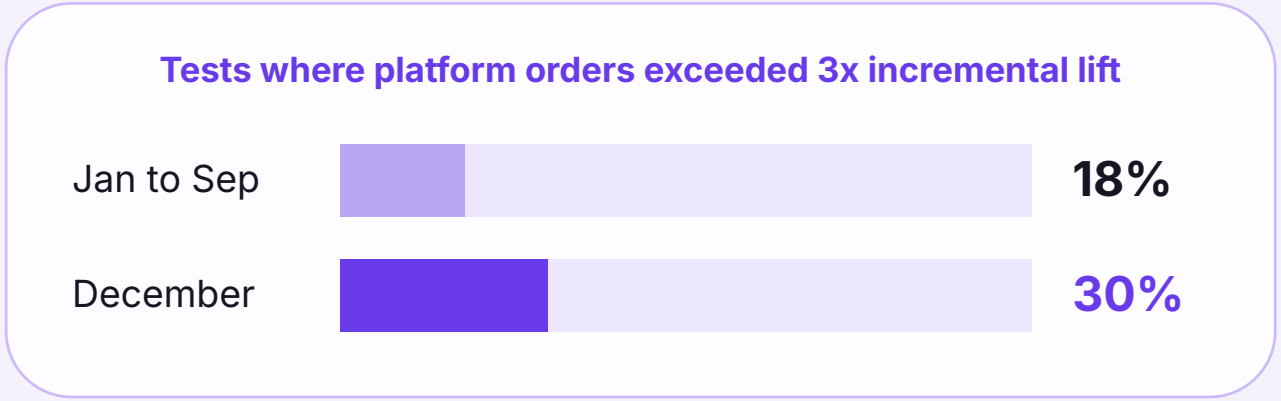
You don’t have to overspend in Q4

Across the same 53 advertisers analyzed, 2025 holiday monthly spend rose 26% and sales rose 28% versus the rest of the year in 2025.

That said, MER dropped by 5%. Brands are right to capitalize on peak seasonality, **but without deeper insight you’re leaving real money on the table.**

When you optimize budgets using insights from lift testing, you can make much better bets **that leverage the holiday demand spike without overspending.**





Copyright © 2026 WorkMagic

Attribution vs. lift

Attribution misses in both directions

Holiday demand can heavily skew attribution. In Q4, platforms credited 1.46 orders for every incremental order. Last-click credited 0.58.

That means that platforms **are reporting a half-order more**, while last-click is **reporting a half-order less**.

That's not all, **30% of December tests** had platform-reported orders **above 3x incremental lift**, versus 18% in January–September.

That's why you have to validate attribution against lift before setting holiday budgets.

SECTION 02

Scoring and Funding the Season



What a good holiday looks like, and how to decide where the next dollar goes.

Q4 2025 test groups by channel type



Primary channels

62%

Google, Meta and TikTok

Emerging channels

38%

AppLovin, CTV and TV, Pinterest, Snapchat and Reddit

Copyright © 2026 WorkMagic

What brands test in Q4

Keeping a pulse on what’s working

Brands kept testing their primary channels through Q4, while more than 1 in 3 tests measured an emerging channel such as AppLovin, CTV or Pinterest.

Testing primary channels with smaller holdouts gives you a sense of **how your main ads perform during peak season**.

Testing emerging channels is a great way **to validate if they’re driving more incremental new customers**, and if they should be scaled for next year.

New Customers is The Key to BFCM, Wherever They Shop

Judge your most important season on incremental new customers, across every storefront.

64%

of incremental orders came from new customers

42%

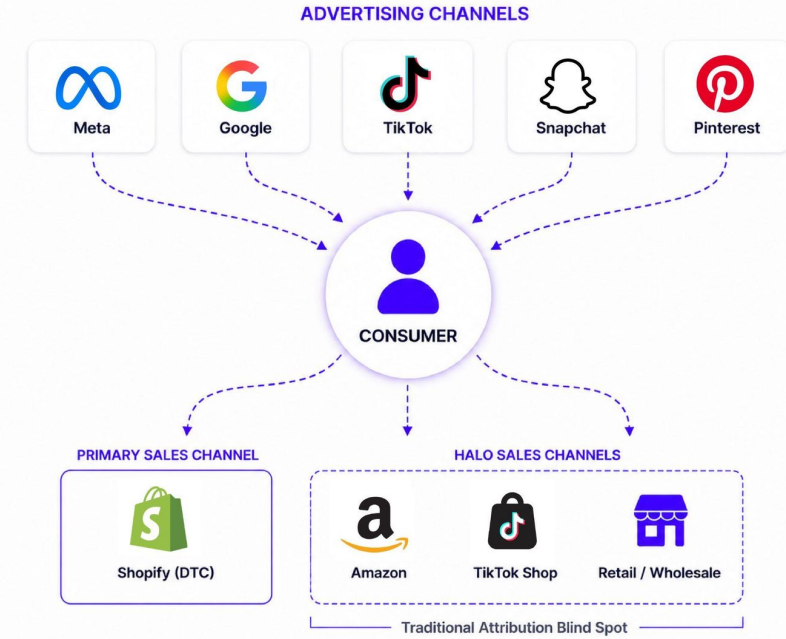
landed off the DTC site, on Amazon, TikTok Shop and retail

Q4 MONTHLY SALES GROWTH VS JAN TO SEP

+49%
Shopify

+31%
Amazon

+58%
TikTok Shop



Copyright © 2026 WorkMagic

Repeat tests on the same brand and media

2025 Q4 vs. January–September 2025

Brand / media	Spend / day	Incr. rev. / day	iROAS, Q1-Q3 → Q4
Brand A · Channel A	+60%	+186%	0.60× → 1.07×
Brand B · Channel B	+63%	+100%	0.30× → 0.37×
Brand C · Channel A	+103%	+33%	3.96× → 2.60×
Brand D · Channel A	+57%	−5%	4.56× → 2.78×
Brand E · TV	+22%	−58%	2.27× → 0.78×
Brand F · CTV	−22%	−64%	1.52× → 0.70×

Copyright © 2026 WorkMagic

Holiday budgets & returns

Higher spend ≠ Higher returns

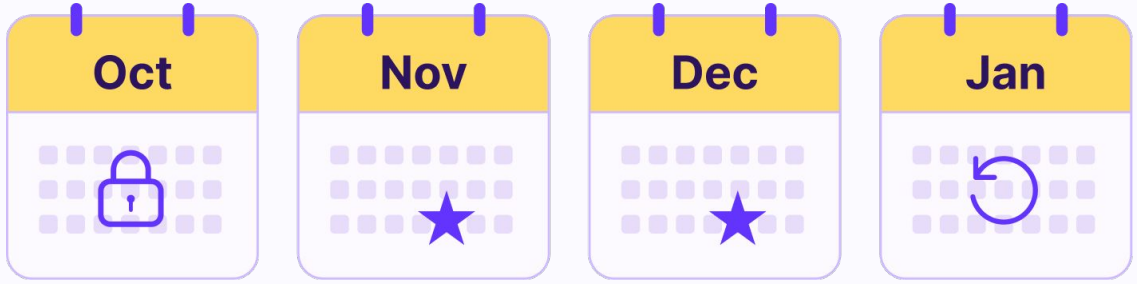
Of five brands with higher Q4 spend per test-day, three grew incremental revenue, but only two improved iROAS.

The same media can deliver different returns when budgets and seasonal demand change. A higher budget needs a new read.

Recheck incremental returns at the holiday budget. Revenue needs to keep pace with spend, and a better iROAS still needs to meet your profitability target.

SECTION 03

Getting Ready For The Holidays

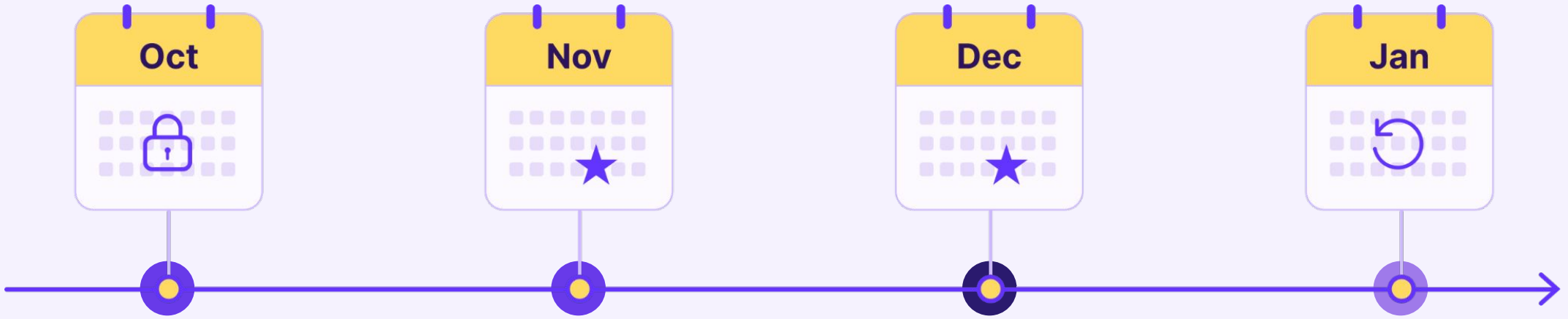


Get your holiday test roadmaps and financial models set up before Q4 gets into full swing.

Calibrate in October, Test Until December, Analyze in January

Brands lock inventory weeks before BFCM. Lock measurement the same way.

BEFORE PEAK · LOCK IT IN | **THROUGH PEAK · KEEP LEARNING**



OCTOBER

Calibrate primary channels

Finish lift tests on your biggest channels so models enter peak on a causal read.

THANKSGIVING

Set goal thresholds

Set iROAS and CAC targets that ground your entire BFCM strategy.

BFCM TO DEC

Schedule tests through peak

Launch-to-measure for new channels, small holdouts where you still need a read.

JANUARY

Read out and retest

Read holiday results, then retest against a business-as-usual baseline.

Copyright © 2026 WorkMagic

Launch-to-measure

SPEND IN SELECTED REGIONS



Test launch

Rest of country stays dark

Pause-to-measure

5%

AS LOW AS

ORDER HOLDOUT



1 in 20 orders

95% remain exposed

Copyright © 2026 WorkMagic

Testing through peak

Run tests without compromising ops

Launch-to-measure adds spend in a few regions while the rest stays dark for that channel.

Pause-to-measure pauses a channel in a few regions. With the right design, a holdout as low as 5% of orders can give a useful read.

Both approaches allow you to gain critical insight without compromising your holiday targets and ad operations.

Q4 testing options

NEW CHANNEL

Launch-to-measure

LARGE EXISTING CHANNEL

Small holdout, or January

SHORT PROMO WINDOW

Test across the window

While a test is live, keep optimizing campaigns as usual.

Copyright © 2026 WorkMagic

Choosing a holiday test

Find a method that fits your use case

New channel: add spend in a small share of regions. Nothing you sell today is held back.

Large existing channel: keep the holdout as small as the design allows, or wait to run a more robust test in January.

Short promotion: run two weeks or more across the window, then read the promo week on its own.

Don’t just test, calibrate your models for Q4 success

Incrementality testing establishes causal truth, calibrating your other models with it allows you to act on it

OCTOBER

Incrementality tests

Establish each channel's causal impact

Calibrated attribution

What should I optimize today?

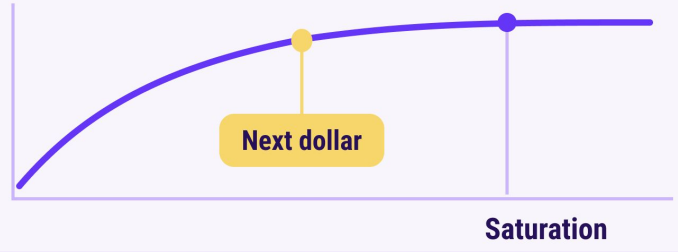
Media Budget Optimizer

Where should the next dollar go?

BEFORE BFCM

Holiday budget split

Set by channel, rechecked as spend scales



Why it matters at peak: Going off of ROAS can’t tell you whether a channel is fully tapped or not, a calibrated MMM readout can.

Curious about how to calibrate your models? Schedule a call with us!

Copyright © 2026 WorkMagic

MYTH #1 “Testing will cost us the season.”	**There’s a test design that fits your risk profile.**
MYTH #2 “The off-season read is good enough.”	**Retest at the holiday budget**
MYTH #3 “A record BFCM means ads worked.”	**Lift tests measure whether they worked**

Holiday testing myths

Testing during Q4 is critical to success

Every test has risk, but you can minimize it. Launch-to-measure holds nothing back; pause-to-measure can read with a holdout as low as 5% of orders.

Retest primary channels every 3–6 months, and recheck at the holiday budget. Incrementality shifts with season and spend.

A record weekend is not proof of ad impact. Holiday demand rises with or without ads. A lift test separates the orders your ads actually drove.

Copyright © 2026 WorkMagic

Key Takeaways

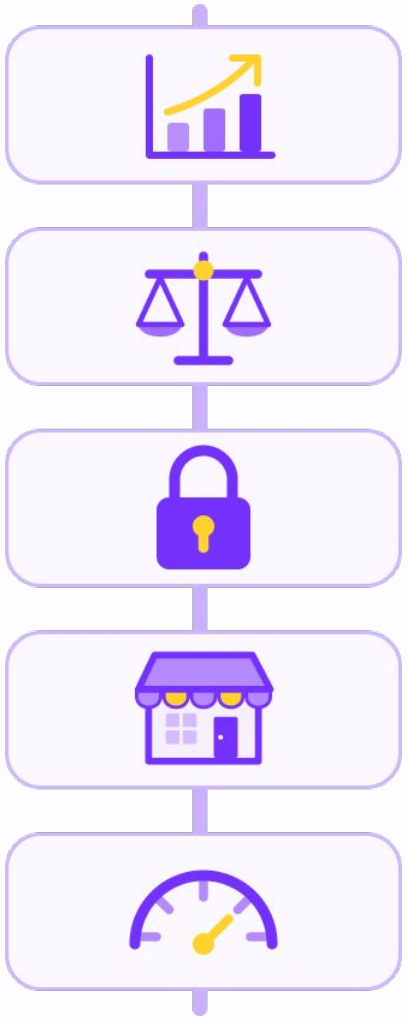
Holiday growth is real, but blended metrics can't explain it. Spend and sales rose together while MER dipped slightly.

Validate attribution against lift before you budget. Our Q4 analysis shows that platforms credited more orders than lift found, and last-click credited fewer.

Lock measurement before peak. Calibrate primary channels in October, then keep testing through BFCM without risking the season.

Score the season on new customers, wherever they buy. Count Amazon, TikTok Shop and retail, not just your site.

Recheck returns at the holiday budget. Bigger budgets produced very different incremental outcomes, even on the same media.



Copyright © 2026 WorkMagic

Most tools just tell you what you sold.

WorkMagic

WorkMagic tells you what your ads actually caused.

Get a Free Measurement Assessment

More resources: The Halo Report | Triangulated Measurement 101 | WorkMagic Case Studies

Copyright © 2026 WorkMagic