

Shopify A/B Testing Decision Sheet

Intelligems vs Shoplift vs VWO vs Optimizely: the seat that runs your test, its speed cost, and the traffic you need. By Kaspian Fuad.

1. Do you have the traffic? (95% conf, 80% power)

2.0% CVR, detect +10% -> ~78,000 sessions per arm
About six weeks for a \$2M/yr store at \$80 AOV.

2.0% CVR, detect +5% -> ~310,000 per arm
Out of reach for most stores under \$5M/yr.

2.0% CVR, detect +20% -> ~19,800 per arm
What a 10,000 per arm, 14-day test can catch.

Under ~\$1M/yr: fix known bugs first, test later
Thin traffic plus a broken baseline = noise.

2. The seat that runs a PRICE test

Intelligems Unlimited \$349/mo (App Store band)
Vendor page: \$1,279/mo up to 200K orders, 3-mo min.

Shoplift Advanced \$399/mo (from \$299)
Priced on monthly unique visitors.

VWO (Wingify) quote only, 30-day trial
AB Tasty merger Jan 2026. No free plan.

Optimizely quote only
"Individually packaged." Enterprise.

3. Entry plans (no price testing)

Intelligems Smart Content \$69/mo
Landing pages, themes, content, personalization.

Intelligems Smart Merchandise \$112.10/mo
Discount types, quantity breaks, bundles, upsells.

Shoplift Core \$99/mo (from \$74)
Template, theme, URL tests. Apply Winner.

4. Pick by the question, not the logo

Price / discount / shipping threshold -> Intelligems
Profit per visitor with your COGS.

Template / theme / landing page -> Shoplift
Launches from the theme customizer.

Multi-site program + analyst -> VWO / Optimizely
Heavier scripts, quote pricing.

5. What the script costs in speed

VWO ~130ms blocking | Optimizely ~160ms
Measured. Library median is 90ms per app.

Intelligems / Shoplift not measured by me yet
Shoplift claims ~2 PSI points. Intelligems: render-block raises FCP/LCP.

Lighthouse flags third-party at 250ms+
INP good = 200ms or under.

6. Anti-flicker settings per tool

Intelligems: element mode > page mode
Hides only price elements. Never async/defer its script.

Shoplift: built-in, decides in <head>
No external request before the decision.

Optimizely: sync snippet, async variations
Sync alone = white screen until loaded.

VWO: never via a tag manager
Tag managers load it async = flicker.

7. Before you launch

Baseline CrUX field data, 28 days
Not one Lighthouse run.

One change per test, one primary metric
RPV or profit per visitor, not CVR alone.

Fix tracking first (GA4, pixels)
A test on broken analytics proves nothing.

8. While it runs

Full weeks only, minimum 14 days
Weekday and payday cycles skew short tests.

No peeking, no early stops
Wait for the sample size, not a green badge.

Watch INP and CLS on the test pages
A slow variant loses for the wrong reason.

9. When it ends (Shopify's own rule)

End the test the day it concludes
shopify.dev: remove the performance cost immediately.

Remove the anti-flicker snippet
"when no tests are running."

Apply the winner in the theme, not the app
Then confirm field data 28 days later.

10. Five-minute verify after install

PageSpeed Insights -> third-party summary
Find the new script and its blocking ms.

Load with JavaScript disabled
Blank page = whole-page anti-flicker is on.

Compare CrUX 28 days later
Lab shows one load. Field shows your shoppers.